

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2006, 09:35:12 ; Search time 109.906 Seconds
(without alignment)
971.459 Million cell updates/sec

Title: US-10-634-108-4
Perfect score: 1303
Sequence: 1 MRPQGPAAAPQLRGLLLLL.....GDASTGNSVSRILIEELPK 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A. Geneseq 21.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1303	100.0	243	4	AAB36667 Human sec
2	1303	100.0	243	5	ABG96340 Human ova
3	1303	100.0	243	5	ABJ05554 Breast ca
4	1303	100.0	243	5	ABB80978 Human REM
5	1303	100.0	243	6	ABR58546 Human can
6	1303	100.0	243	6	ABR48227 Human bla
7	1303	100.0	243	6	ABG75758 Human REM
8	1303	100.0	243	6	ABU56607 Lung canc
9	1303	100.0	243	7	ADB80510 Ovarian c
10	1303	100.0	243	7	ADN39855 Cancer/an
11	1303	100.0	243	7	ADN38732 Cancer/an
12	1303	100.0	243	8	ADL70254 LBFU301 p
13	1303	100.0	243	8	ADU23513 Alternati
14	1303	100.0	243	8	ADU06493 Novel bro
15	1303	100.0	243	9	AEb28819 Human CYP
16	1303	100.0	243	3	AAB08856 Amino aci
17	1303	100.0	278	3	AAm25746 Human pro
18	1303	100.0	278	5	ABG96338 Human ova
19	1303	100.0	278	6	ABR47627 Breast ca
20	1303	100.0	278	9	ADZ51345 Amino aci
21	1300	99.8	243	3	AAy91529 Human sec
22	1300	99.8	243	8	ADL71601 Novel hum
23	1300	99.8	278	5	ABP68631 Human pan
24	1298	99.6	243	6	ABJ37031 Human bre

ALIGNMENTS

RESULT 1
AAB36667
ID AAB36667 standard; protein; 243 AA.

AC AAB36667;

DT 14-MAR-2001 (first entry)

DE Human secretory protein TGC-628 SEQ ID NO:7.

KW Human; secretory protein; cancer; immune disease; infectious disease;
KW lung function disorder; liver function disorder; antiinflammatory;
KW gastrointestinal disorder; cytostatic; haematopoietic; anticoagulant;
KW immunomodulatory; hepatotropic; cell proliferation-stimulant;
KW cell migratory agent; cell differentiation-inducer.

OS Homo sapiens.

XX WO2000071581-A1.

XX 30-NOV-2000.

XX 19-MAY-2000; 2000WO-JP003221.

XX 20-MAY-1999; 99JP-00140229.

XX (TAKE) TAKEDA CHEM IND LTD.

PI Itoh Y, Mogi S, Tanaka H, Ohkubo S, Ogi K;

DR WPI; 2001-032023/04.

XX N-PSDB; AAC90707.

XX Novel secretory protein and its salt with e.g. anti-cancer, anti-inflammatory and hematopoietic, effects, applicable as drugs in remedies and preventives to treat diseases like cancer and immune diseases.

XX Claim 1; Page 89-90; 122pp (Japanese).

XX AAC90701 to AAC90715 encode the human secretory proteins given in
XX AAB36661 to AAB36675. The proteins can have cytostatic, anti-inflammatory, haematopoietic, anti-coagulant, immunomodulatory and hepatotropic activities, and can be used as cell migratory agents, cell proliferation-stimulants and cell differentiation-inducers. The proteins are useful in the treatment and prevention of diseases such as cancer, lung function disorder, liver function disorder, gastrointestinal disorder and immune diseases. AAC90716 to AAC90755 represent PCR primers

10/19/2006

CC which are used in the exemplification of the present invention
XX
SQ Sequence 243 AA;

Query Match 100.0%; Score 1303; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 1e-120;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRPGPAASPORLRLGLLLLLLQLPAPSSASEIPKQKQKQALRQREVVDLYNGMCLQGPA 60
Db 1 MRPGPAASPORLRLGLLLLLLQLPAPSSASEIPKQKQKQALRQREVVDLYNGMCLQGPA 60

Qy 61 GVPGRDGSPPGANGIPGTGIPGRDGFKEGECLESFEESWTNPKQCSWSSLYNYGIDL 120
Db 61 GVPGRDGSPPGANGIPGTGIPGRDGFKEGECLESFEESWTNPKQCSWSSLYNYGIDL 120

Qy 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACQWYFTFNGAECSPLEIAIYLDQ 180
Db 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACQWYFTFNGAECSPLEIAIYLDQ 180

Qy 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGMNSVSRITIEE 240
Db 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGMNSVSRITIEE 240

Qy 241 LPK 243
Db 241 LPK 243

RESULT 2
ABG96340
ID ABG96340 standard; protein; 243 AA.

XX AC ABG96340;
XX
XX 11-DEC-2002 (first entry)
XX
XX Human ovarian cancer marker M450.

Human; ovarian cancer; marker; cancer; familial history; brain disorder;
central nervous system disorder; bacterial meningitis; viral meningitis;
Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
brain herniation; inflammation; encephalitis; testicular disorder;
nontuberculous granulomatous orchitis; connective tissue disorder;
heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
histological type; carcinogenic; ovarian cancer marker.

OS Homo sapiens.
XX
XX WO200271928-A2.
XX
XX 19-SEP-2002.
XX
XX 14-MAR-2002; 2002WO-US007826.
XX
XX 14-MAR-2001; 2001US-0276025P.
XX
XX 14-MAR-2001; 2001US-0276026P.
XX
XX 10-AUG-2001; 2001US-0311732P.
XX
XX 19-SEP-2001; 2001US-0323580P.
XX
XX 26-SEP-2001; 2001US-0324967P.
XX
XX 26-SEP-2001; 2001US-0325102P.
XX
XX 26-SEP-2001; 2001US-0325149P.

(MILL-) MILLENNIUM PHARM INC.

XX Monahan JE, Gannavarapu M, Hoerach S, Kamatkar S, Kovatis SG;
XX MeyerB RE, Morrissey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
XX Baat RC, Lu K, Schmandt RE, Zhao X, Glatt K;
XX
XX WPI; 2002-723277/78.
XX
XX N-PSDB; ABS76436.

Assessing whether a patient is afflicted with ovarian cancer, useful in

PT assessing the stage or progression of the disease, comprises comparing
PT the expression level of a cancer marker in a sample from a patient and
PT from a non cancer patient.

XX Disclosure; Page 252; 481pp; English.

XX The present invention relates to a new method for assessing whether a
XX patient is afflicted with ovarian cancer. The method involves comparing
XX the expression level of a marker in a patient sample and the normal level
XX of expression of the marker in a control non-ovarian cancer sample, where
XX the marker is selected from 363 cancer markers described in the
XX specification. The method of the invention is useful in diagnosing or
XX characterising cancer, in detecting the presence of cancer as early as
XX possible, and the recurrence of ovarian cancer. The method may also be of
XX particular use with patients having an enhanced risk of developing
XX ovarian cancer (e.g. patients having a familial history of ovarian
XX cancer). The cancer markers may be used in the management and treatment
XX of e.g. brain and central nervous system disorders (e.g. bacterial and
XX viral meningitis, Alzheimer's disease or Parkinson's disease), brain
XX disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
XX inflammations (e.g. bacterial or viral meningitis or encephalitis),
XX testicular disorders (e.g. nontuberculous granulomatous orchitis),
XX connective tissue disorders, or heart disorders (e.g. ischaemic heart
XX disease or atherosclerosis). The compositions and methods may also be
XX used in assessing the histological type of neoplasm associated with
XX ovarian cancer, monitoring the progression of ovarian cancer, determining
XX whether ovarian cancer has metastasized or is likely to metastasize,
XX selecting a composition for inhibiting ovarian cancer, assessing the
XX ovarian carcinogenic potential of a compound, or inhibiting ovarian
XX cancer or at risk of developing ovarian cancer. The present amino acid
XX sequence represents one of the ovarian cancer markers described in the
XX invention

SQ Sequence 243 AA;

Query Match 100.0%; Score 1303; DB 5; Length 243;
Best Local Similarity 100.0%; Pred. No. 1e-120;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRPGPAASPORLRLGLLLLLLQLPAPSSASEIPKQKQKQALRQREVVDLYNGMCLQGPA 60
Db 1 MRPGPAASPORLRLGLLLLLLQLPAPSSASEIPKQKQKQALRQREVVDLYNGMCLQGPA 60

Qy 61 GVPGRDGSPPGANGIPGTGIPGRDGFKEGECLESFEESWTNPKQCSWSSLYNYGIDL 120
Db 61 GVPGRDGSPPGANGIPGTGIPGRDGFKEGECLESFEESWTNPKQCSWSSLYNYGIDL 120

Qy 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACQWYFTFNGAECSPLEIAIYLDQ 180
Db 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACQWYFTFNGAECSPLEIAIYLDQ 180

Qy 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGMNSVSRITIEE 240
Db 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGMNSVSRITIEE 240

Qy 241 LPK 243
Db 241 LPK 243

RESULT 3
ABJ05554

ID ABJ05554 standard; protein; 243 AA.

XX AC ABJ05554;

XX 14-NOV-2002 (first entry)

XX Breast cancer-associated protein 19.

XX Breast cancer; breast cancer-associated gene sequence; drug development;
XX pharmacogenetics; biosensor development.

OS Unidentified.
 XX WO200259377-A2.
 PN 01-AUG-2002.
 PD 24-JAN-2002; 2002WO-US002242.
 PF 24-JAN-2001; 2001US-026396SP.
 XX 02-FEB-2001; 2001US-0265928P.
 PR 09-APR-2001; 2001US-00829472.
 PR 09-APR-2001; 2001US-0282698P.
 PR 04-MAY-2001; 2001US-0288590P.
 PR 29-MAY-2001; 2001US-0294443P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 PA Mack DH, Gish KC, Afar D;
 XX WPI; 2002-583738/62.
 XX N-PSDB; ABT07711.
 DR Detecting a breast cancer-associated transcript in a patient's cell,
 PT useful for diagnosing breast cancer, comprises contacting a biological
 PT sample with a polynucleotide that selectively hybridizes with breast
 PT cancer nucleic acids.
 XX Disclosure; Page 364; 414pp; English.
 PS The invention comprises a method of detecting a breast cancer-associated
 CC transcript in a cell from a patient. The method of the invention involves
 CC contacting a biological sample from the patient with a nucleotide that
 CC hybridizes to one of the 69 breast cancer-associated gene sequences shown
 CC in the specification. The method of the invention is useful in the
 CC diagnosis or prognosis of breast cancer, and for detecting genes that are
 CC up or down-regulated in breast cancer cells. Genes identified by the
 CC method of the invention can be used in diagnostic purposes and also as
 CC targets for screening for therapeutic compounds that modulate breast
 CC cancer (e.g. hormones or antibodies). Identification of genes that are
 CC over or under expressed in breast cancer can additionally provide high-
 CC resolution, high-sensitivity datasets which can be used in the areas of
 CC diagnostics, therapeutics, drug development, pharmacogenetics, protein
 CC structure and biosensor development. Amino acid sequences AB05536 -
 CC AB05604 represent the proteins encoded by the 69 breast cancer-
 CC associated genes of the invention
 XX Sequence 243 AA;

Query Match 100.0%; Score 1303; DB 5; Length 243;
 Best Local Similarity 100.0%; Pred. No. 1e-120;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPQGAASPORLGLLLLLQLPAPSSASEIPKGKQAKQLRQREVVDLYNMCLOGPA 60
 DB 1 MRPQGAASPORLGLLLLLQLPAPSSASEIPKGKQAKQLRQREVVDLYNMCLOGPA 60
 QY 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGECLRESFEESWTPNYKQCSWSSLYGIDL 120
 DB 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGECLRESFEESWTPNYKQCSWSSLYGIDL 120
 QY 121 GKIAECTTKMRSNALSRLVFGSRLKCRNACQWTFNFGACSGPLPIEAIYLDQ 180
 DB 121 GKIAECTTKMRSNALSRLVFGSRLKCRNACQWTFNFGACSGPLPIEAIYLDQ 180
 QY 181 GSPENNSTINHTSSVEGLCEGAGLVDAIIVGTCSDYPKGDASTGWSVRIIEE 240
 DB 181 GSPENNSTINHTSSVEGLCEGAGLVDAIIVGTCSDYPKGDASTGWSVRIIEE 240
 QY 241 LPK 243
 DB 241 LPK 243

RESULT 4
 ABB80978
 ID ABB80978 standard; protein; 243 AA.
 XX ABB80978;
 AC 21-OCT-2002 (first entry)
 DT Human REMODELIN polypeptide.
 DE REMODELIN; vulnery; vasotrophic; cytostatic; osteopathic; collagen;
 KW gene therapy; bone; human.
 XX Homo sapiens.
 OS WO200242487-A2.
 PN 30-MAY-2002.
 PD 19-OCT-2001; 2001WO-US050940.
 XX 19-OCT-2000; 2000US-00692081.
 PF (MAIN-) MAINE MEDICAL CENT RES INST.
 PR Linder V, Friesel R;
 XX WPI; 2002-590472/63.
 DR N-PSDB; ABN86481.
 XX New mammalian REMODELIN polypeptide for diagnosing arterial restenosis,
 PT negative remodeling, fibrosis, collagen disease, and bone disease in a
 PT mammal.
 XX Claim 5; Fig 4B; 187pp; English.
 PS The invention relates to mammalian REMODELIN polypeptides and encoding
 CC polynucleotides. REMODELIN is used to diagnose arterial restenosis,
 CC negative remodeling or fibrosis, bone disease such as osteogenesis
 CC imperfecta (OI), collagen disease such as OI, dystrophic epidermolysis
 CC bullosa (DEB) and Bethlem myopathy in a mammal. The polynucleotides are
 CC used to affect cellular gene expression in a mammal, where the cellular
 CC gene is from transforming growth factor (TGF)-beta1, collagen IIIalpha1,
 CC osteopontin, biglycan, alkaline phosphatase or bone morphogenic protein
 CC 4. The expression of osteopontin is dependent on Cbfa1. Compositions
 CC comprising antisense REMODELIN sequences are useful for treating diseases
 CC mediated by abnormal expression of a REMODELIN molecule in a human such
 CC as impaired wound healing, fibrosis of an organ, ectopic ossification, or
 CC hypertrophic scar formation. REMODELIN is useful in the development of
 CC bone during mammalian embryogenesis, proliferation and/or migration,
 CC cellular signaling, adventitial fibrosis, negative remodeling and arterial
 CC restenosis, smooth muscle cell proliferation, and arterial remodeling and
 CC useful in gene therapy. REMODELIN is useful for elucidating the function
 CC of REMODELIN molecules in a cell, to identify a compound that affects
 CC REMODELIN expression and/or TGF-beta signaling, as a potential
 CC therapeutic drug candidate for arterial restenosis, anti-cancer therapy,
 CC to promote or inhibit wound healing, to inhibit scar tissue or keloid
 CC formation, to promote bone fracture healing, and to increase or decrease
 CC expression of REMODELIN in mammal. It is also useful for producing
 CC recombinant cell and transgenic non-human mammals which are useful tools
 CC for the study of REMODELIN action, for identifying novel diagnostics and
 CC therapeutics for treatment, and for elucidating the cellular roles of
 CC REMODELIN. The present sequence represents a human REMODELIN polypeptide
 XX Sequence 243 AA;

Query Match 100.0%; Score 1303; DB 5; Length 243;
 Best Local Similarity 100.0%; Pred. No. 1e-120;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPQGAASPORLGLLLLLQLPAPSSASEIPKGKQAKQLRQREVVDLYNMCLOGPA 60
 DB 1 MRPQGAASPORLGLLLLLQLPAPSSASEIPKGKQAKQLRQREVVDLYNMCLOGPA 60

XX The present invention describes a method for detecting a bladder cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with a polynucleotide
CC that selectively hybridizes to a sequence that is 80 % identical to a
CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC encode the human bladder cancer-associated proteins given in ABR48146 to
CC ABR48242). Bladder cancer-associated sequences from the present invention
CC have cytostatic activities, and can be used in antisense gene therapy and
CC in vaccine production. The method can be used for detecting a bladder
CC cancer-associated transcript in a cell from a patient. The method is
CC useful in diagnosing or treating bladder cancer and in screening for
CC compounds that modulate bladder cancer, such as hormones or antibodies.
CC The nucleic acid molecules from the present invention may be used in
CC various screening and diagnostic methods, and for gene therapy, vaccine
CC and/or antisense/inhibition applications

XX Sequence 243 AA;

Query Match 100.0%; Score 1303; DB 6; Length 243;

Best Local Similarity 100.0%; Pred. No. 1e-120; Mismatches 0; Indels 0; Gaps 0;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPQGPAAAPQRLRGLLLLLLLLQLPAPSSASIPKQKQALRQREVVDLYNGMCLQGPA 60

Db 1 MRPQGPAAAPQRLRGLLLLLLLLQLPAPSSASIPKQKQALRQREVVDLYNGMCLQGPA 60

QY 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESPEESWTPNYKQCSWSLNYGIDL 120

Db 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESPEESWTPNYKQCSWSLNYGIDL 120

QY 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACORWYFTFNGAECGSLPIEAIYLDQ 180

Db 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACORWYFTFNGAECGSLPIEAIYLDQ 180

QY 181 GSPENNSTINHTSSVEGLCEGIGAGLVDAIWDVGTCDYKPGDASTGWNVSRIIEE 240

Db 181 GSPENNSTINHTSSVEGLCEGIGAGLVDAIWDVGTCDYKPGDASTGWNVSRIIEE 240

QY 241 LPK 243

Db 241 LPK 243

Query Match 100.0%; Score 1303; DB 6; Length 243;

Best Local Similarity 100.0%; Pred. No. 1e-120; Mismatches 0; Indels 0; Gaps 0;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPQGPAAAPQRLRGLLLLLLLLQLPAPSSASIPKQKQALRQREVVDLYNGMCLQGPA 60

Db 1 MRPQGPAAAPQRLRGLLLLLLLLQLPAPSSASIPKQKQALRQREVVDLYNGMCLQGPA 60

QY 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESPEESWTPNYKQCSWSLNYGIDL 120

Db 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESPEESWTPNYKQCSWSLNYGIDL 120

QY 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACORWYFTFNGAECGSLPIEAIYLDQ 180

Db 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACORWYFTFNGAECGSLPIEAIYLDQ 180

QY 181 GSPENNSTINHTSSVEGLCEGIGAGLVDAIWDVGTCDYKPGDASTGWNVSRIIEE 240

Db 181 GSPENNSTINHTSSVEGLCEGIGAGLVDAIWDVGTCDYKPGDASTGWNVSRIIEE 240

QY 241 LPK 243

Db 241 LPK 243

Query Match 100.0%; Score 1303; DB 6; Length 243;

Best Local Similarity 100.0%; Pred. No. 1e-120; Mismatches 0; Indels 0; Gaps 0;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPQGPAAAPQRLRGLLLLLLLLQLPAPSSASIPKQKQALRQREVVDLYNGMCLQGPA 60

Db 1 MRPQGPAAAPQRLRGLLLLLLLLQLPAPSSASIPKQKQALRQREVVDLYNGMCLQGPA 60

QY 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESPEESWTPNYKQCSWSLNYGIDL 120

Db 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESPEESWTPNYKQCSWSLNYGIDL 120

QY 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACORWYFTFNGAECGSLPIEAIYLDQ 180

Db 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACORWYFTFNGAECGSLPIEAIYLDQ 180

QY 181 GSPENNSTINHTSSVEGLCEGIGAGLVDAIWDVGTCDYKPGDASTGWNVSRIIEE 240

Db 181 GSPENNSTINHTSSVEGLCEGIGAGLVDAIWDVGTCDYKPGDASTGWNVSRIIEE 240

QY 241 LPK 243

Db 241 LPK 243

Query Match 100.0%; Score 1303; DB 6; Length 243;

Best Local Similarity 100.0%; Pred. No. 1e-120; Mismatches 0; Indels 0; Gaps 0;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPQGPAAAPQRLRGLLLLLLLLQLPAPSSASIPKQKQALRQREVVDLYNGMCLQGPA 60

Db 1 MRPQGPAAAPQRLRGLLLLLLLLQLPAPSSASIPKQKQALRQREVVDLYNGMCLQGPA 60

QY 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESPEESWTPNYKQCSWSLNYGIDL 120

Db 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESPEESWTPNYKQCSWSLNYGIDL 120

XX

PA

PI

XX

XX

DR

DR

XX

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

(LINDNER V.

(FRIE//) FRIESEL R E.

Lindner V, Friesel RE;

WPI; 2003-238238/23.

N-PSDB; ABX11341.

New isolated REMODELIN nucleic acid and polypeptide, useful for mediating

arterial remodeling, formation of bone and cartilage, and the diagnosis

and treatment of disorders associated with aberrant expression of

REMODELIN.

Claim 3; Fig 4B; 81pp; English.

The invention discloses an isolated nucleic acid encoding a mammalian

adventitia-inducible bone expressed molecule called REMODELIN (or REMODEL

and/or adventitia induced bone expressed molecule, AIBE). REMODELIN

expression was induced by transforming growth factor beta (TGF-beta),

which is important because proliferative events occurring in the

adventitia contribute to vascular remodelling and restenosis in the

to vascular injury and TGF-beta has been shown to be a factor involved in

this. Also disclosed is an antibody raised against REMODELIN and methods

for treating a disease mediated by abnormal expression of a REMODELIN in

a human, for identifying a compound that affects or reduces expression of

REMODELIN in a cell (e.g. antisense therapy), for identifying a compound

that affects TGF-beta signalling and for increasing or reducing REMODELIN

expression in a mammal, comprising administering a REMODELIN expression

increasing or reducing TGF-beta to the mammal, thereby increasing

REMODELIN expression or inhibiting signalling via the TGF-beta receptor

type II and reducing expression of REMODELIN in the mammal. The methods

and compositions of the present invention are useful for mediating

arterial remodelling, formation of bone and cartilage and the diagnosis

and treatment of disorders associated with aberrant expression of

REMODELIN, such as osteogenesis imperfecta, dystrophic epidermolysis

bullosa, Bethlem myopathy, negative remodelling, wound healing, arterial

stenosis, vessel injury, fibrosis and calcification of a transplant,

preferably a heart valve transplant. The sequence presented is the human

REMODELIN protein

Sequence 243 AA;

Query Match 100.0%; Score 1303; DB 6; Length 243;

Best Local Similarity 100.0%; Pred. No. 1e-120;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPQGPAAAPQRLRGLLLLLLLLQLPAPSSASIPKQKQALRQREVVDLYNGMCLQGPA 60

Db 1 MRPQGPAAAPQRLRGLLLLLLLLQLPAPSSASIPKQKQALRQREVVDLYNGMCLQGPA 60

QY 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESPEESWTPNYKQCSWSLNYGIDL 120

Db 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESPEESWTPNYKQCSWSLNYGIDL 120

QY 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACORWYFTFNGAECGSLPIEAIYLDQ 180

Db 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACORWYFTFNGAECGSLPIEAIYLDQ 180

QY 181 GSPENNSTINHTSSVEGLCEGIGAGLVDAIWDVGTCDYKPGDASTGWNVSRIIEE 240

Db 181 GSPENNSTINHTSSVEGLCEGIGAGLVDAIWDVGTCDYKPGDASTGWNVSRIIEE 240

QY 241 LPK 243

Db 241 LPK 243

Query Match 100.0%; Score 1303; DB 6; Length 243;

Best Local Similarity 100.0%; Pred. No. 1e-120; Mismatches 0; Indels 0; Gaps 0;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPQGPAAAPQRLRGLLLLLLLLQLPAPSSASIPKQKQALRQREVVDLYNGMCLQGPA 60

Db 1 MRPQGPAAAPQRLRGLLLLLLLLQLPAPSSASIPKQKQALRQREVVDLYNGMCLQGPA 60

QY 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESPEESWTPNYKQCSWSLNYGIDL 120

Db 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESPEESWTPNYKQCSWSLNYGIDL 120

QY 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACORWYFTFNGAECGSLPIEAIYLDQ 180

Db 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACORWYFTFNGAECGSLPIEAIYLDQ 180

QY 181 GSPENNSTINHTSSVEGLCEGIGAGLVDAIWDVGTCDYKPGDASTGWNVSRIIEE 240

Db 181 GSPENNSTINHTSSVEGLCEGIGAGLVDAIWDVGTCDYKPGDASTGWNVSRIIEE 240

QY 241 LPK 243

Db 241 LPK 243

Query Match 100.0%; Score 1303; DB 6; Length 243;

Best Local Similarity 100.0%; Pred. No. 1e-120; Mismatches 0; Indels 0; Gaps 0;

RESULT 8

ABU56607

ID ABU56607 standard; protein; 243 AA.

XX

AC ABU56607;

```
XX 02-APR-2003 (first entry)
XX Lung cancer-associated polypeptide #200.
XX Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX Unidentified.
XX OS
XX WO200286443-A2.
XX 31-OCT-2002.
XX 18-APR-2002; 2002WO-US012476.
XX 18-APR-2001; 2001US-0284770P.
XX 10-MAY-2001; 2001US-0290492P.
XX 09-NOV-2001; 2001US-0339245P.
XX 13-NOV-2001; 2001US-0350666P.
XX 29-NOV-2001; 2001US-0334370P.
XX 12-APR-2002; 2002US-0372246P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX Aziz N, Murray R;
PI WPI; 2003-093161/08.
DR N-PSDB; ABX76336.
XX
XX Detecting a lung cancer-associated transcript in a cell from a patient
XX for treating lung cancer, by contacting a biological sample from the
XX patient with a polynucleotide that exhibits increased or decreased
XX expression in lung cancer.
XX Claim 27; Page 337; 453pp; English.
XX The invention relates to a method for detecting a lung cancer-associated
XX transcript in a cell from a patient, comprising contacting a biological
XX sample from the patient with a polynucleotide that selectively hybridizes
XX to a sequence that is at least 80 % identical to a gene that exhibits
XX increased or decreased expression in lung cancer samples. Lung cancer-
XX associated polynucleotides and polypeptides are used for identifying a
XX compound that modulates a lung cancer-associated polypeptide, for
XX inhibiting proliferation of a lung cancer-associated cell to treat lung
XX cancer in a patient and for treating a mammal having lung cancer by
XX administering a modulatory compound identified. The methods are useful
XX for treating lung cancer, such as small cell lung cancer, non-small cell
XX lung cancer or other benign or precancerous lesions, e.g. atelectasis,
XX emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
XX hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
XX bronchiectasis. The genes, polynucleotides and polypeptides are useful
XX for diagnostic purposes and as targets for screening for therapeutic
XX compounds that modulate lung cancer, such as antibodies. Sequences
XX ABU56408-ABU56745 represent lung cancer-associated polypeptides of the
XX invention
XX Sequence 243 AA;
XX Query Match 100.0%; Score 1303; DB 6; Length 243;
XX Best Local Similarity 100.0%; Pred. No. 1e-120;
XX Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX 1 MRPQGPAPSPQLRGLLLLLLLOLPAPSASEIPKGGKQALRQREVDVLYNGMCLQGPA 60
XX DB 1 MRPQGPAPSPQLRGLLLLLLLOLPAPSASEIPKGGKQALRQREVDVLYNGMCLQGPA 60
XX 61 GVPGRDGSFGANGIGTGPICRQDGFKEGKGCLESPFEESWTPNYKQCSWSSLYGIDL 120
XX DB 61 GVPGRDGSFGANGIGTGPICRQDGFKEGKGCLESPFEESWTPNYKQCSWSSLYGIDL 120
Qy 121 GKIAECTTKMRSNSALRVLPSSGLRLKCRNACCORWYFTENGAECSGLPIEAIYLDQ 180
Db 121 GKIAECTTKMRSNSALRVLPSSGLRLKCRNACCORWYFTENGAECSGLPIEAIYLDQ 180
Qy 181 GSPMNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRIIIEE 240
Db 181 GSPMNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRIIIEE 240
Qy 241 LPK 243
Db 241 LPK 243
RESULT 9
ADB80510
ID ADB80510 standard; protein; 243 AA.
XX ADB80510;
XX 04-DEC-2003 (first entry)
XX Ovarian cancer-associated protein #37.
XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
XX post-operative chemotherapy; radiation therapy; tumour prognosis;
XX pre-cancerous lesion detection.
XX Homo sapiens.
XX WO2002102235-A2.
XX 27-DEC-2002.
XX 18-JUN-2002; 2002WO-US019297.
XX 18-JUN-2001; 2001US-0299234P.
XX 27-AUG-2001; 2001US-0315287P.
XX 05-SEP-2001; 2001US-0317544P.
XX 13-NOV-2001; 2001US-0350666P.
XX 12-APR-2002; 2002US-0372246P.
XX (EOSB-) EOS BIOTECHNOLOGY INC;
XX Mack DH, Gish KC;
XX WPI; 2003-167431/16.
XX N-PSDB; ADB80509.
XX Detecting an ovarian cancer-associated transcript in a cell from a
XX patient, comprises contacting a biological sample from the patient with a
XX polynucleotide that hybridizes to an ovarian cancer gene.
XX Claim 13; Page 299; 332pp; English.
XX The invention relates to a method of detecting an ovarian cancer-
XX associated transcript in a cell from a patient, by contacting a
XX biological sample from the patient with a polynucleotide that selectively
XX hybridizes to a sequence at least 80% identical to any of one of 80
XX nucleic acid sequences given in the specification. The method is useful
XX in diagnosing ovarian cancer and in identifying and using agents and/or
XX targets that inhibit ovarian cancer. The nucleic acid molecule,
XX polypeptide and the antibody may also be used in detecting ovarian
XX cancers, monitoring and early detection of relapse following treatment,
XX monitoring response to therapy, selecting patients for post-operative
XX chemotherapy or radiation therapy, in selecting mode of therapy,
XX determining tumour prognosis, early detection of pre-cancerous lesions,
XX and as vaccines. This sequence corresponds to one of the proteins used
XX for the detection method of the invention.
XX Sequence 243 AA;
XX Query Match 100.0%; Score 1303; DB 7; Length 243;
```

Best Local Similarity 100.0%; Pred. No. 1e-120; Mismatches 0; Indels 0; Gaps 0;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRPQGPAAAPQRLGRLGLLLLLQLPAPSSASIPKQKQAKLRQREVVDLYNGMCLQGPA 60
Db 1 MRPQGPAAAPQRLGRLGLLLLLQLPAPSSASIPKQKQAKLRQREVVDLYNGMCLQGPA 60
QY 61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKGLRESPEESWTNPKQCSWSLNYGIDL 120
Db 61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKGLRESPEESWTNPKQCSWSLNYGIDL 120
QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACORWYFTFNGAECSGPLPIEAIYLDQ 180
Db 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACORWYFTFNGAECSGPLPIEAIYLDQ 180
QY 181 GSPENNSTINIHTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGWNVSRIIEE 240
Db 181 GSPENNSTINIHTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGWNVSRIIEE 240
QY 241 LPK 243
Db 241 LPK 243

RESULT 10
ADN39855
ID ADN39855 standard; protein; 243 AA.
AC ADN39855;
DT 17-JUN-2004 (first entry)
XX Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C225.
DE Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
vulnerary; gene therapy; vaccine.
XX Homo sapiens.
OS
XX WO2003042661-A2.
PN
XX 22-MAY-2003.
PD
XX 13-NOV-2002; 2002WO-US036810.
PF
XX 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397757P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
PA
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

XX WPI: 2003-468649/44.
DR N-PSDB; ADN39637.
XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX Claim 12; SEQ ID NO C225; 1385pp; English.
PS The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
XX whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX Sequence 243 AA;
SQ Query Match 100.0%; Score 1303; DB 7; Length 243;
Best Local Similarity 100.0%; Pred. No. 1e-120; Mismatches 0; Indels 0; Gaps 0;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRPQGPAAAPQRLGRLGLLLLLQLPAPSSASIPKQKQAKLRQREVVDLYNGMCLQGPA 60
Db 1 MRPQGPAAAPQRLGRLGLLLLLQLPAPSSASIPKQKQAKLRQREVVDLYNGMCLQGPA 60
QY 61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKGLRESPEESWTNPKQCSWSLNYGIDL 120
Db 61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKGLRESPEESWTNPKQCSWSLNYGIDL 120
QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACORWYFTFNGAECSGPLPIEAIYLDQ 180
Db 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACORWYFTFNGAECSGPLPIEAIYLDQ 180
QY 181 GSPENNSTINIHTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGWNVSRIIEE 240
Db 181 GSPENNSTINIHTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGWNVSRIIEE 240
QY 241 LPK 243
Db 241 LPK 243
RESULT 11
ADN38732
ID ADN38732 standard; protein; 243 AA.
XX AC ADN38732;
XX 17-JUN-2004 (first entry)
DT
XX Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:50.
DE Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
vulnerary; gene therapy; vaccine.
XX

CC neoplasm occurring in soft tissue, bone, breast, cervix, colon, endometrium, esophagus, kidney, larynx, liver, lung, omentum, ovary, pancreas, rectum, thyroid, myometrium, prostate, skin, small intestine, bladder, spleen or stomach.

XX Sequence 243 AA;

Query Match 100.0%; Score 1303; DB 8; Length 243;
Best Local Similarity 100.0%; Pred. No. 1e-120;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPQGPAAAPQRLRGLLLLLQLPAPSSASEIPKQKQAKLRQREVVDLYNGMCLQGPA 60
DB 1 MRPQGPAAAPQRLRGLLLLLQLPAPSSASEIPKQKQAKLRQREVVDLYNGMCLQGPA 60
QY 61 GVPGRDGPANGIPGTPGIPGRDGFKEGKGECLRESFEESWTPNYKQCSWSLNYGIDL 120
DB 61 GVPGRDGPANGIPGTPGIPGRDGFKEGKGECLRESFEESWTPNYKQCSWSLNYGIDL 120
QY 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIYLDQ 180
DB 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIYLDQ 180
QY 181 GSPENNSTINIHRSTSSVEGLCEGIGAGLVDAIWWGTCSDYPKGDASTGWNVSRIIEE 240
DB 181 GSPENNSTINIHRSTSSVEGLCEGIGAGLVDAIWWGTCSDYPKGDASTGWNVSRIIEE 240
QY 241 LPK 243
DB 241 LPK 243

RESULT 13
ADU23513
ID ADU23513 standard; protein; 243 AA.

XX AC ADU23513;

XX DT 27-JAN-2005 (first entry)

XX DE Alternative human STOP-1 protein SeqID 3.

XX KW STOP-1; tumour; angiogenesis; vasculogenesis; cytostatic; gene therapy.

XX OS Homo sapiens.

XX PN WO2004094476-A2.

XX PD 04-NOV-2004.

XX PF 16-APR-2004; 2004WO-US011793.

XX PR 16-APR-2003; 2003US-0463656P.

XX PA (GETH) GENENTECH INC.

XX PI Ackerly H, Ashkenazi A, Eberhard D, Frantz G, French D, Fuh G;

XX PI Hongo J, Lee C, Marsters S, Pitti R, Raab H, Soroceanu L;

XX PI Varfolomeev E, Wolf B;

XX DR WPI; 2004-795534/78.

XX PT New monoclonal antibody that specifically binds to an oligomeric form of human STOP-1, useful for diagnosing, preventing or treating diseases involving angiogenesis and vasculogenesis, e.g. cancer, and in research or drug screening.

XX PS Example 1; SEQ ID NO 3; 265pp; English.

XX CC This invention relates to a novel monoclonal antibody that specifically binds to an oligomeric form of human STOP-1. Specifically, it refers to a STOP-1 polypeptide variant that cannot be secreted from a cell or form a disulfide bind with another STOP-1 protein. The present invention

CC describes a composition comprising a STOP-1 antagonist and a pharmaceutical carrier, where the antagonist specifically binds STOP-1 and the binding can be inhibited by the monoclonal antibody, 6B12. As such, it provides methods for diagnosing or monitoring a tumour in a patient, as well as methods for the prevention or inhibition of the tumour growth that overexpresses STOP-1. Furthermore, it provides a drug screening method for testing the activity of a candidate antagonist or agonist of STOP-1 such that it can be used to treat a disease or condition associated with excessive, inappropriate or uncontrolled angiogenesis and/or vasculogenesis in a mammalian subject, alternatively it can also induce angiogenesis in a patient that would benefit from increased angiogenesis. Accordingly, pharmaceutical compositions developed thereof exhibit cytostatic activities and can be used in gene therapy as angiogenesis stimulators or inhibitors. This polypeptide sequence is an alternative human STOP-1 protein of the invention.

XX Sequence 243 AA;

Query Match 100.0%; Score 1303; DB 8; Length 243;
Best Local Similarity 100.0%; Pred. No. 1e-120;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPQGPAAAPQRLRGLLLLLQLPAPSSASEIPKQKQAKLRQREVVDLYNGMCLQGPA 60
DB 1 MRPQGPAAAPQRLRGLLLLLQLPAPSSASEIPKQKQAKLRQREVVDLYNGMCLQGPA 60
QY 61 GVPGRDGPANGIPGTPGIPGRDGFKEGKGECLRESFEESWTPNYKQCSWSLNYGIDL 120
DB 61 GVPGRDGPANGIPGTPGIPGRDGFKEGKGECLRESFEESWTPNYKQCSWSLNYGIDL 120
QY 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIYLDQ 180
DB 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIYLDQ 180
QY 181 GSPENNSTINIHRSTSSVEGLCEGIGAGLVDAIWWGTCSDYPKGDASTGWNVSRIIEE 240
DB 181 GSPENNSTINIHRSTSSVEGLCEGIGAGLVDAIWWGTCSDYPKGDASTGWNVSRIIEE 240
QY 241 LPK 243
DB 241 LPK 243

RESULT 14

ADU06493

ID ADU06493 standard; protein; 243 AA.

XX AC ADU06493;

XX DT 27-JAN-2005 (first entry)

XX DE Novel bronchial cancer-associated human protein SeqID717.

XX KW bronchial cancer; cytostatic; tumour-associated protein;

XX KW cancer detection; metastasis; tumour; human.

XX OS Homo sapiens.

XX PN DE10316701-A1.

XX PD 04-NOV-2004.

XX PF 09-APR-2003; 2003DE-01016701.

XX PR 09-APR-2003; 2003DE-01016701.

XX PA (HINZ/) HINZMANN B.

XX PA (HERM/) HERMANN K.

XX PA (CAST/) HEIDEN CASTANOS-VELEZ E.

XX PI Mennerich D, Bruemendorf T, Heiden E, Hinzmann B, Rosenthal A, Pillarsky C;

XX PI Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pillarsky C;

DR WPI; 2004-786403/78.
 DR N-PSDB; ADU06006.
 XX New nucleic acid, and derived proteins, useful for diagnosis of bronchial
 PT cancer and in screening for therapeutic and diagnostic agents.
 XX
 PS Claim 2; SEQ ID NO 717; 1381pp; German.
 XX
 CC This invention relates to a novel isolated nucleic acid associated with
 CC bronchial cancer comprising 489 defined sequences given in the
 CC specification. The invention may be useful for the production of
 CC compounds with a cytostatic activity through the inhibition of expression
 CC or activity of tumour-associated proteins. The novel DNA sequences and
 CC the proteins/peptides encoded by them are used for detecting bronchial
 CC cancer or determining the risk of developing it and to screen for
 CC specific binding partners of the DNA or protein sequences, where the
 CC binding partners are potentially useful as agents for treating or
 CC diagnosing bronchial cancer. The DNA or protein sequences can also be
 CC used for prognosis, detection of metastases and for secondary treatment
 CC (of tumours that have been stabilised or are no longer detectable).
 CC Detecting abnormal expression of the DNA sequences provides early
 CC diagnosis of bronchial cancers. The present sequence is that of a protein
 CC encoded by a novel bronchial cancer-associated human gene sequence of the
 CC invention.
 XX
 SQ Sequence 243 AA;

Query Match 100.0%; Score 1303; DB 8; Length 243;
 Best Local Similarity 100.0%; Pred. No. 1e-120;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRQGPAPASPORLRLGLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 60
 DB 1 MRQGPAPASPORLRLGLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 60
 QY 61 GVPGRDGGPGANGIPGTGPIGRDGFKEGKECLRESFEESWTNPKYKCSWSSLYGIDL 120
 DB 61 GVPGRDGGPGANGIPGTGPIGRDGFKEGKECLRESFEESWTNPKYKCSWSSLYGIDL 120
 QY 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACQORWYFTFNGAECSPLEIATYLDQ 180
 DB 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACQORWYFTFNGAECSPLEIATYLDQ 180
 QY 181 GSPMNSTINHTSSVEGLCEGICAGLVDVAIWGTCTSDYPKGDASTGWSVSRIIIEE 240
 DB 181 GSPMNSTINHTSSVEGLCEGICAGLVDVAIWGTCTSDYPKGDASTGWSVSRIIIEE 240
 QY 241 LPK 243
 DB 241 LPK 243

RESULT 15
 AEB28819 standard; protein; 243 AA.
 XX AEB28819;
 AC AEB28819;
 XX 08-SEP-2005 (first entry)
 DT Human CTHRC1 protein SEQ ID NO:4.
 XX collagen triple helix repeat containing 1; CTHRC1; vulnery;
 KW antiinflammatory; respiratory-gen.; vasotropic; fibrosis; restenosis;
 XX wound healing; scarring.
 XX Homo sapiens.
 XX US2005147602-A1.
 XX 07-JUL-2005.
 XX 10-SEP-2004; 2004US-00939233.

XX 19-OCT-2000; 2000US-00692081.
 PR 19-OCT-2001; 2001US-00045992.
 PR 18-SEP-2003; 2003US-0503933P.
 PR 18-SEP-2003; 2003US-0504107P.
 XX (MAIN-) MAINE MEDICAL CENT RES INST.
 XX Lindner V;
 PI WPI; 2005-478076/48.
 DR N-PSDB; AEB28834.
 XX Novel isolated polypeptide comprising human cleaved collagen triple helix
 PT repeat containing 1 (CTHRC1) or isolated mutant CTHRC1 polypeptide,
 PT useful for treating or preventing disease mediated by collagen matrix
 PT production e.g. fibrosis.
 XX
 PS Example 1; SEQ ID NO 4; 115pp; English.

CC The invention relates to an isolated polypeptide (I) comprising a human
 CC cleaved collagen triple helix repeat containing 1 (CTHRC1) or an isolated
 CC mutant CTHRC1 polypeptide (II) comprising substitution of a human CTHRC1
 CC collagen domain with a mouse collagen 1 alpha 1 collagen domain. (I) and
 CC (II) are useful for treating or preventing a disease mediated by collagen
 CC matrix production in a human, which involves administering to a human
 CC afflicted with the disease an effective amount of CTHRC1, where the
 CC diseases chosen from fibrosis, constrictive remodeling and restenosis.
 CC (I) or (II) is useful for decreasing levels of bone morphogenetic protein
 CC 1 (BMP1) or BMP1 mRNA in a cell, increasing the level of a propeptide
 CC (chosen from procollagen and a propeptide of lysyl-oxidase) in a cell,
 CC inhibiting collagen formation by a cell, decreasing bone matrix formation
 CC by a cell, decreasing the level of collagen in a cell, increasing the
 CC level of procollagen in a cell, decreasing collagen formation in a mammal
 CC having a condition mediated by collagen formation, where the condition is
 CC chosen from wound scarring, wound healing, keloid formation, inflammation
 CC -associated scarring, pulmonary fibrosis, and angioplasty-associated
 CC vascular fibrosis, or increasing the level of chordin in a cell, which
 CC involves contacting the cell with (I) or (II). (I) is useful for
 CC inhibiting cross-linking of collagen fibrils in a cell, which involves
 CC contacting a cell with a BMP1 inhibiting amount of (I) or (II), where
 CC BMP1 is responsible for processing a propeptide lysyl-oxidase, and
 CC further where the lysyl-oxidase mediates cross-linking of the collagen
 CC fibrils, thus inhibiting cross-linking of collagen fibrils in the cell.
 CC (I) is useful for treating a disease mediated by expression of BMP1 in a
 CC mammal, increasing the level of bone morphogenetic protein 4 (BMP4) in a
 CC cell, increasing the level of bone morphogenetic protein 4 (BMP4) in a
 CC bone growth in a mammal, promoting differentiation of a stem cell,
 CC decreasing the level of osteopontin (OPN) in a cell, treating a disease
 CC mediated by under expression of BMP4 in a mammal in need, and increasing
 CC the level of a muscle segment homeobox 1 (Mex1) in a cell. The present
 CC sequence represents human CTHRC1, which is used in the exemplification of
 CC the present invention.

Sequence 243 AA;

Query Match 100.0%; Score 1303; DB 9; Length 243;
 Best Local Similarity 100.0%; Pred. No. 1e-120;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRQGPAPASPORLRLGLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 60
 DB 1 MRQGPAPASPORLRLGLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 60
 QY 61 GVPGRDGGPGANGIPGTGPIGRDGFKEGKECLRESFEESWTNPKYKCSWSSLYGIDL 120
 DB 61 GVPGRDGGPGANGIPGTGPIGRDGFKEGKECLRESFEESWTNPKYKCSWSSLYGIDL 120
 QY 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACQORWYFTFNGAECSPLEIATYLDQ 180
 DB 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACQORWYFTFNGAECSPLEIATYLDQ 180
 QY 181 GSPMNSTINHTSSVEGLCEGICAGLVDVAIWGTCTSDYPKGDASTGWSVSRIIIEE 240

Db 181 GSPFNNSTINIHRTSSVEGLCEGIGAGLVDAIHWGTCSDDYKGDASTGWNVSRIIEE 240
QY 241 LFK 243
Db 241 LFK 243

Search completed: April 20, 2006, 09:41:16
Job time : 112.906 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (C) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model
Run on: April 20, 2006, 09:41:38 ; Search time 19.6941 Seconds
(without alignments)
1187.191 Million cell updates/sec

Title: US-10-634-108-4
Perfect score: 1303
Sequence: 1 MRQGPAAAPQRLRGLLLLL.....GDASTGNSVSRILIIIEELPK 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	131.5	10.1	246	2 S29328	complement subcomp
2	129	9.9	1752	2 A45407	collagen alpha 3(I)
3	128	9.8	289	2 T20177	hypothetical prote
4	127.5	9.8	1464	1 CGHU1S	collagen alpha 1(I)
5	126	9.7	327	2 T29031	hypothetical prote
6	126	9.7	754	2 A55267	collagen alpha 5(I)
7	125	9.6	888	2 S28791	collagen alpha 1(X)
8	125	9.6	1670	1 CGHU3B	collagen alpha 3(I)
9	124	9.5	1466	1 CGHU7L	collagen alpha 1(I)
10	124	9.5	1763	2 S16366	collagen alpha 2(I)
11	123.5	9.5	311	2 T15268	hypothetical prote
12	123	9.4	304	2 T26185	hypothetical prote
13	123	9.4	304	2 T26184	hypothetical prote
14	123	9.4	1691	1 S22917	collagen alpha 5(I)
15	122.5	9.4	300	2 T24482	hypothetical prote
16	122.5	9.4	920	2 A45748	collagen alpha 1(V)
17	122.5	9.4	1549	2 T48103	type VII collagen
18	122	9.4	2944	2 A54849	collagen alpha 1(V)
19	121	9.3	178	2 A39762	collagen alpha 1(X)
20	121	9.3	325	2 T18594	hypothetical prote
21	120.5	9.2	428	2 T24769	hypothetical prote
22	120	9.2	245	1 C1HUQC	complement subcomp
23	120	9.2	358	2 T26281	hypothetical prote
24	119.5	9.2	298	2 T27644	hypothetical prote
25	119.5	9.2	310	2 T29731	hypothetical prote
26	119.5	9.2	458	2 T31631	hypothetical prote
27	119.5	9.2	671	1 CGBT1S	collagen alpha 1(I)
28	119.5	9.2	1453	2 S21626	collagen alpha 1(I)
29	119.5	9.2	1669	1 CGHU4B	collagen alpha 1(I)

30	119	9.1	298	2 JCI448	collagen col-34 -
31	119	9.1	299	2 T29956	hypothetical prote
32	119	9.1	1042	1 CGCH1S	collagen alpha 1(I)
33	119	9.1	1049	1 CGBO7S	collagen alpha 1(I)
34	119	9.1	1806	1 CGHU1E	collagen alpha 1(X)
35	118.5	9.1	177	2 S37749	collagen alpha 2(X)
36	118.5	9.1	290	2 T24586	hypothetical prote
37	118.5	9.1	330	2 S46657	collagen alpha 1(X)
38	118.5	9.1	488	2 A27353	collagen alpha 1(I)
39	118.5	9.1	1691	1 CGHU6B	collagen alpha 6(I)
40	118	9.1	1669	1 CGMS4B	collagen alpha 1(I)
41	118	9.1	1744	2 S40991	collagen alpha 1(I)
42	117.5	9.0	323	2 A61396	collagen alpha 1(I)
43	117.5	9.0	1486	1 B40333	collagen alpha 1(I)
44	117.5	9.0	1492	2 A40333	collagen alpha 1(I)
45	117	9.0	248	1 LNHUP1	pulmonary surfacta

ALIGNMENTS

RESULT 1

S29328
complement subcomponent Clq chain C - mouse
C:Species: Mus musculus (house mouse)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S29328
R:Petry, F.; Reid, K.B.M.; Loos, M.
Eur. J. Biochem. 209, 129-134, 1992
A>Title: Isolation, sequence analysis and characterization of cDNA clones coding for the ecerbellin.
A:Reference number: S29328; MUID:93011118; PMID:1396691
A:Accession: S29328
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-246 <PEP>
A:Cross-references: UNIPROT:Q02105; UNIPARC:UPI0000028DD0; EMBL:X66295; NID:950228; PIDD:122222
C:Superfamily: complement subcomponent Clq chain A; complement Clq carboxyl-terminal homology F:122-245/Domain: complement Clq carboxyl-terminal homology <Clq>

Query Match	10.1%	Score 131.5;	DB 2;	Length 246;
Best Local Similarity	38.5%	Pred. No. 0.00025;		
Matches	35;	Conservative	5;	Mismatches 32; Indels 19; Gaps 2;
QY	5	GPAAAPQRLRGLLLLLLQLPAPASASRIPKQKQAKLRQREVVDLYNMCICLOGPAGVPG	64	
Db	4	GPSCQPQCGLCLLLFLALLPLRSQAS-----AGCYGIPGMPGMPGAPG	47	
QY	65	RDGS---PGANGIPGTGPIGPDGPKGKGE	92	
Db	48	KDGHDLQGPKEGPIPAVPGTQGPKGKGE	78	

RESULT 2

A45407
collagen alpha 3(IV) chain - sea urchin (Strongylocentrotus purpuratus)
C:Species: Strongylocentrotus purpuratus (purple urchin)
C>Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 31-Dec-2004
C:Accession: A45407; A43903; A23940
R:Exposito, J.Y.; D'Alessio, M.; Di Liberto, M.; Ramirez, F.
J. Biol. Chem. 268, 5249-5254, 1993
A>Title: Complete primary structure of a sea urchin type IV collagen alpha chain and an A:Reference number: A45407; MUID:93186842; PMID:8444899
A:Accession: A45407
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-1752 <EXP>
A:Cross-references: UNIPROT:Q26312; UNIPARC:UPI00001773DE
A>Note: sequence extracted from NCBI backbone (NCBIP:126841)
R:Wessel, G.M.; Etkin, M.; Benson, S.
Dev. Biol. 148, 261-272, 1991
A>Title: Primary mesenchyme cells of the sea urchin embryo require an autonomously prod
A:Reference number: A43903; MUID:92038439; PMID:1936564

```
A:Accession: A43903
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 'P',633-1537,'G' <WS>
A:Cross-references: UNIPARC:UPI000007C802; GB:S64572; NID:q238616; PIDN:AAB20270.1; PID:
A:Note: sequence extracted from NCBI backbone (NCBIN:64572, NCBI:P:64573)
R:Venkatesan, M.; De Pablo, F.; Vogeli, G.; Simpson, R.T.
Proc. Natl. Acad. Sci. U.S.A. 83, 3351-3355, 1986
A:Title: Structure and developmentally regulated expression of a Strongylocentrotus purpurus
A:Reference number: A23940; MUID:86205894; PMID:3458186
A:Accession: A23940
A:Molecule type: DNA
A:Residues: 742-812 <VEN>
A:Cross-references: UNIPARC:UPI00001773DF; EMBL:M13206
A:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix
F:129-161/Domain: amino-terminal nonhelical, 78 <7SD>
F:162-1523/Region: interrupted helical
F:1524-1752/Domain: carboxyl-terminal nonhelical, NCI <NC1>
F:1534-1634/Domain: collagen IV carboxyl-terminal repeat <CT1>
F:1644-1748/Domain: collagen IV carboxyl-terminal repeat <CT2>
F:129/Modified site: allysine (Lys) #status Predicted
Query Match 9.9%; Score 129; DB 2; Length 1752;
Best Local Similarity 22.9%; Pred. No. 0.0038;
Matches 47; Conservative 20; Mismatches 60; Indels 78; Gaps 9;
QY 3 PQGPAASPORLURLLLLLLLLQLPAPSSASEIPKGKQAKLRQREVVDLYNGMCLQGPGAGV 62
Db 1453 PQGPGRDGR-----PGPQG----PPG-----LTGDKGT 1477
QY 63 PGRGSPGANGIPGTPGPRDGFGEKGECLRESF-----EESWT 103
Db 1478 PGVQNGPVGSGVPGFPGGLKGBQGFQNGQPGDPGFGTKGEAIPGSSGFFFTIRHSQT 1537
QY 104 PNYKQC-----SW-----SSLYNGIDLKIAECTFTKRSNSALRVLFSGSLRL 147
Db 1538 TSIPQCPQGTAKWHGYSLLFVQGNERGHQDLKFGSC----LKRFSTMFLFC-NINN 1592
QY 148 KCRNACCORWYTFNGAECSPGLPI 172
Db 1593 VCHVASRNDYSWLSLTSTB---PMPM 1614
RESULT 3
T20177
hypothetical protein C53B4.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20177
R:Berkas, M.
submitted to the EMBL Data Library, December 1995
A:Reference number: Z19233
A:Accession: T20177
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-289 <WIL>
A:Cross-references: UNIPROT:Q18799; UNIPARC:UPI0000080F8D; EMBL:Z68215; PIDN:CAA92453.1;
A:Experimental source: clone C53B4
C:Genetics:
A:Gene: CESP:C53B4.5
A:Map position: 4
Query Match 9.8%; Score 128; DB 2; Length 289;
Best Local Similarity 35.7%; Pred. No. 0.00061;
Matches 30; Conservative 5; Mismatches 19; Indels 30; Gaps 2;
QY 25 PAPSSASIPKGKQAKLRQREVVDLYNGMCLQGAGVPGRDGS-----P 69
Db 203 PQGPSPGFGPRGPQGS-----RGPAGQFGKDGAGQGGPKGANGEPGQP 247
QY 70 GANGIPGTPGTPGPRDGFGEKGECC 93
Db 248 GRDQGPFGPQGRDGHFGEKGVC 271
```

```
RESULT 4
CGHUIS
collagen alpha 1(I) chain precursor - human
N:Alternate names: procollagen alpha 1(I) chain
C:Species: Homo sapiens (man)
C:Date: 12-Aug-1981 #sequence_revision 04-Oct-1996 #text_change 31-Dec-2004
A:Accession: T60114; S01143; A93335; I55254; A39943; I55237; A35233; S09400; B90567; S11
5269; A29439; I53466; A02852; I37247
R:D'Alessio, M.; Bernard, M.; Pretorius, P.J.; de Wet, W.; Ramirez, F.; Pretorius, P.J.
Gene 67, 105-115, 1988
A:Title: Complete nucleotide sequence of the region encompassing the first twenty-five e
A:Reference number: T60114; MUID:88329734; PMID:2843432
A:Accession: T60114
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-369,'L',371-589 <DAL>
A:Cross-references: UNIPROT:P02452; UNIPROT:Q14992; UNIPROT:Q16053; UNIPROT:Q13896; UNIP
R:Tromp, G.; Kuivaniemi, H.; Stacey, A.; Shikata, H.; Baldwin, C.T.; Jaenisch, R.; Prock
Biochem. J. 253, 919-922, 1988
A:Title: Structure of a full-length cDNA clone for the prepro-alpha-1(I) chain of human
A:Reference number: S01143; MUID:89025644; PMID:3178743
A:Accession: S01143
A:Molecule type: mRNA
A:Residues: 1-472 <TRO>
A:Cross-references: UNIPARC:UPI000016A6F9; EMBL:X07884; NID:g30015; PIDN:CAA30731.1; PID
A:Note: submitted to the EMBL/GenBank/DBDJ databases by Prockop, D.J., 13-JUN-1988
R:Chu, M.L.; de Wet, W.; Bernard, M.; Ding, J.F.; Morabito, M.; Myers, J.; Williams, C.;
Nature 310, 337-340, 1984
A:Title: Human proalpha1(I) collagen gene structure reveals evolutionary conservation of
A:Reference number: A93335; MUID:84270697; PMID:6462220
A:Accession: A93335
A:Molecule type: DNA
A:Residues: 1-58,'Q',60-181 <CHU>
A:Cross-references: UNIPARC:UPI0000173B3C; EMBL:X00820; NID:g35657; PIDN:CAA25394.1; PID
R:Roosouw, C.M.S.; Vargers, W.P.; du Plooy, S.J.; Bernard, M.P.; Ramirez, F.; de Wet, W.
J. Biol. Chem. 262, 15151-15157, 1987
A:Title: DNA sequences in the first intron of the human pro-alpha 1(I) collagen gene en
A:Reference number: I55254; MUID:88033098; PMID:2822714
A:Accession: I55254
A:Status: translation not shown; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-45 <ROS>
A:Cross-references: UNIPARC:UPI000016A6B2; GB:J02829; NID:g180387; PIDN:AAA51993.1; PID:
R:Bornstein, P.; McKay, J.; Morishima, J.K.; Devarayalu, S.; Gellinas, R.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 8869-8873, 1987
A:Title: Regulatory elements in the first intron contribute to transcriptional control o
A:Reference number: A39943; MUID:88097389; PMID:3480516
A:Accession: A39943
A:Molecule type: DNA
A:Residues: 1-34 <BOR>
A:Cross-references: UNIPARC:UPI000016A6B1; GB:J03559; NID:g180876; PIDN:AAA52052.1; PID:
R:Chu, M.L.; de Wet, W.; Bernard, M.; Ramirez, F.
J. Biol. Chem. 260, 2315-2320, 1985
A:Title: Fine structural analysis of the human pro-alpha 1 (I) collagen gene. Promoter s
A:Reference number: I55237; MUID:85130970; PMID:2857713
A:Accession: I55237
A:Status: translation not shown; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-34 <CH2>
A:Cross-references: UNIPARC:UPI000016A6B1; GB:M10627; NID:g180383; PIDN:AAA51992.1; PID:
R:Wirtz, M.K.; Keene, D.R.; Hori, H.; Glanville, R.W.; Steinmann, B.; Rao, V.H.; Hollist
J. Biol. Chem. 265, 6312-6317, 1990
A:Title: In vivo and in vitro noncovalent association of excised alpha1(I) amino-termina
tome, type VII.
A:Reference number: A35233; MUID:90202908; PMID:2318855
A:Accession: A35233
A:Molecule type: protein
A:Residues: 33-52 <WIR>
A:Cross-references: UNIPARC:UPI0000173B3D
A:Note: this propeptide fragment remained non-covalently bound to a defective, uncleaved
R:Weil, D.; d'Alessio, M.; Ramirez, F.; de Wet, W.; Cole, W.G.; Chan, D.; Bateman, J.F.
```



```
Db      386 GNPAGDQPGKANGAPCIAGPFGARG 416

RESULT 5
T29031
hypothetical protein F53G12.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29031
R:Wu, X.; Graves, T.
submitted to the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid F53G12.
A:Reference number: Z20555
A:Accession: T29031
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-327 <WUX>
A:Cross-references: UNIPROT:O01799; UNIPARC:UPI0000078PFF; EMBL:AF003139; PIDN:AAB54156.
A:Experimental source: strain Bristol N2; clone F53G12
C:Genetics:
A:Gene: CESP:F53G12.7
A:Map position: 1
A:Introns: 59/3; 138/1; 223/2

Query Match      9.7%; Score 126; DB 2; Length 327;
Best Local Similarity 34.4%; Pred. No. 0.0011;
Matches 32; Conservative 2; Mismatches 21; Indels 38; Gaps 3;

Qy      1 MRPGPAASQRLRGLLLLLLQLPAPSSASBIPKQKQKQALRQREVVDLYNGMCLQGPA 60
Db      261 MCPQP-----PQPG-----PQKQGA-----GGEK 282

Qy      61 GVPGRDGGPGANGIPGTGIPGRDGFKEGKGC 93
Db      283 GAGPDQGNPGYKPGQPGAPGPDGSGAGEKGC 315

RESULT 6
A55267
collagen alpha 5(IV) chain - dog (fragment)
C:Species: Canis lupus familiaris (dog)
C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C:Accession: A55267
R:Zheng, K.; Thorner, P.S.; Marrano, P.; Bauml, R.; McInnes, R.R.
Proc. Natl. Acad. Sci. U.S.A. 91, 3989-3993, 1994
A:Title: Canine X chromosome-linked hereditary nephritis: a genetic model for human X-linked type IV.
A:Reference number: A55267; MUID:94224868; PMID:8171024
A:Accession: A55267
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-754 <ZHE>
A:Cross-references: UNIPROT:Q28247; UNIPARC:UPI0000126D56; GB:U07888; NID:5469547; PIDN:
C:Superfamily: collagen alpha 1(IV) chain

Query Match      9.7%; Score 126; DB 2; Length 754;
Best Local Similarity 35.4%; Pred. No. 0.0027;
Matches 34; Conservative 7; Mismatches 33; Indels 22; Gaps 4;

Qy      3 PQGPAASQRLRGLLLLLLQLPAPSSASEIPKQKQKQALRQREVVDLYNGMC 55
Db      332 PQGPPRP-----GLPGEGRGUPGNIGKGRGNPGQPGQGLPLGLKGD-- 377

Qy      56 LQGPAGVPRDGGSPGANGIPGTGIPGRDGFKEGKGC 91
Db      378 -QGPGIQGNPRGPLNGMKGDGGLPGVPGPGMGK 412

RESULT 7
S28791
collagen alpha 1(XI) chain - chicken (fragment)
C:Species: Gallus gallus (chicken)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
```

```
C:Accession: S28791
R:Nah, H.D.; Barembaum, M.; Upholt, W.B.
J. Biol. Chem. 267, 22581-22586, 1992
A:Title: The chicken alpha1(XI) collagen gene is widely expressed in embryonic tissues.
A:Reference number: S28791; MUID:93054557; PMID:1429607
A:Accession: S28791
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-888 <NAH>
A:Cross-references: UNIPROT:Q90796; UNIPARC:UPI00000FBAD3; EMBL:M88593; NID:G211619; PID:
C:Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homology
P:665-887/Domain: fibrillar collagen carboxyl-terminal homology <PCC>

Query Match      9.6%; Score 125; DB 2; Length 888;
Best Local Similarity 35.2%; Pred. No. 0.004;
Matches 32; Conservative 4; Mismatches 19; Indels 36; Gaps 3;

Qy      3 PQGPAA--SPQRLGRLLLLLLQLPAPSSASEIPKQKQKQALRQREVVDLYNGMCLQGPA 60
Db      479 PQGPAGKPGPEGLRGI-----PQVGE-----QSLP 504

Qy      61 GVPGRDGGPGANGIPGTGIPGRDGFKEGKGC 91
Db      505 GAGQDGGPGHLGPGGLPGLKGDGPGSGKEGC 535

RESULT 8
CGHU3B
collagen alpha 3(IV) chain precursor, long splice form - human
N:Alternate names: Goodpasture antigen; procollagen alpha 3(IV) chain long splice form
C:Species: Homo sapiens (man)
C>Date: 28-Oct-1994 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C:Accession: A54763; A43928; A44043; A45971; A39786
R:Mariyama, M.; Leinonen, A.; Mochizuki, T.; Tryggvason, K.; Reiders, S.T.
J. Biol. Chem. 269, 23013-23017, 1994
A:Title: Complete primary structure of the human alpha3(IV) collagen chain. Coexpression
A:Reference number: A54763; MUID:94364994; PMID:8083201
A:Accession: A54763
A:Molecule type: mRNA
A:Residues: 1-1670 <MAR>
A:Cross-references: UNIPROT:Q01955; UNIPARC:UPI0000173BE3; GB:X80031; NID:G577563; PID:G
A:Experimental source: kidney
R:Turner, N.; Mason, P.J.; Brown, R.; Fox, M.; Povey, S.; Rees, A.; Pusey, C.D.
J. Clin. Invest. 89, 592-601, 1992
A:Title: Molecular cloning of the human Goodpasture antigen demonstrates it to be the al
A:Reference number: A43928; MUID:92147878; PMID:1737849
A:Accession: A43928
A:Molecule type: mRNA
A:Residues: 1331-1524, 'I', 1526-1670 <TUR>
A:Cross-references: UNIPARC:UPI0000173BE4; GB:M81379
A:Experimental source: kidney
R:Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.
J. Biol. Chem. 267, 19780-19784, 1992
A:Title: Exon/intron structure of the human alpha 3(IV) gene encompassing the Goodpastur
ction.
A:Reference number: A44043; MUID:93015826; PMID:1400291
A:Accession: A44043
A:Molecule type: DNA; mRNA
A:Residues: 1386-1670 <QUI>
A:Cross-references: UNIPARC:UPI000016A42D; GB:M92993; NID:G177895; PIDN:AAA21610.1; PID:
A>Note: sequence extracted from NCHI backbone (NCBIP:115597)
R:Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.
J. Biol. Chem. 269, 17358, 1994
A:Reference number: A44738; MUID:94274734; PMID:8006044
A:Contents: annotation; erratum; correction to intronic sequence in A44043
R:Bernal, D.; Quinones, S.; Saus, J.
J. Biol. Chem. 268, 12090-12094, 1993
A:Title: The human mRNA encoding the Goodpasture antigen is alternatively spliced.
A:Reference number: A45971; MUID:93280184; PMID:8505332
A:Accession: A45971
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1427-1444 <BER>
```


A:Cross-references: UNIPARC:UPI0000173BES
A:Note: sequence extracted from NCBI backbone (NCBIP:133363); sequence incorrectly identified
R:Morrisson, K.E.; Mariyama, M.; Yang-Feng, T.L.; Reenders, S.T.
Am. J. Hum. Genet. 49, 545-554, 1991
A:Title: Sequence and localization of a partial cDNA encoding the human alpha3 chain of
A:Reference number: A39786; MUID:9133570; PMID:1882840
A:Accession: A39786
A:Molecule type: mRNA
A:Residues: 1453-1593, 'A', 1595-1670 <MOR>
A:Cross-references: UNIPARC:UPI000014C40B; GB:S55790; NID:G234418; PIDN:AAB19637.1; PID:
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit
ed and subsequently O-glycosylated.
C:Comment: In Goodpasture's syndrome, an autoimmune response develops against an epitope
C:Genetics:
A:Gene: GDB:COL4A3
A:Cross-references: GDB:128351; OMIM:120070
A:Map position: 2q36-2q37
A:Introns: 1385/1; 1418/1; 1488/1; 1547/2; 1585/3; 1643/2 #status incomplete
A:Note: the alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands with
C:Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha 3
among trimer amino-terminal domains (with disulfide and desmosine cross-links), dimeric
er associations in the interrupted helical domain (with disulfide and desmosine cross-li
C:Function:
A:Description: minor structural component of extracellular basement membrane in kidney
C:Superfamily: collagen alpha 1(IV) chain
C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extracel
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-1670/Product: collagen alpha 3(IV) chain, long splice form #status predicted <MAT>
F:29-42/Domain: amino-terminal nonhelical, NH1 <NH1>
F:43-1438/Region: interrupted helical
F:791-793/Region: cell attachment (R-G-D) motif
F:996-998/Region: cell attachment (R-G-D) motif
F:1154-1156/Region: cell attachment (R-G-D) motif
F:1306-1308/Region: cell attachment (R-G-D) motif
F:1345-1347/Region: cell attachment (R-G-D) motif
F:1432-1434/Region: cell attachment (R-G-D) motif
F:1439-1670/Domain: carboxyl-terminal nonhelical, NC1 <NC1>
F:1451-1551/Domain: collagen IV carboxyl-terminal repeat <CRL>
F:1561-1665/Domain: collagen IV carboxyl-terminal repeat <CR2>
F:31,33,39,41,125,422,476,479,682,722,809,1387/Disulfide bonds: interchain #status predi
F:253/Binding site: carboxylate (Asn) (covalent) #status predicted
F:1460-1548,1493-1551/disulfide bonds: (or 1460-1551, 1493-1548) #status predicted
F:1505-1511,1616-1622/disulfide bonds: #status predicted
F:1570-1662,1604-1665/disulfide bonds: (or 1570-1665, 1604-1662) #status predicted
Query Match 9.6%; Score 125; DB 1; Length 1670;
Best Local Similarity 36.2%; Pred. No. 0.0083;
Matches 34; Conservative 6; Mismatches 36; Indels 18; Gaps 3;
QY 5 GPAASPORLRGLLLLLLQLLPAPSSASEIPKQKQALRQREVVDLYNGMCLQGPGAVPG 64
DB 467 GPKGEP-----GLLCTQCPVPGPGPLGLPLGLHGVKGIPGR-----QGAAGLKG 511
QY 65 RDGSPGANGIPGTGIPGRD---GFKGEGECLR 95
DB 512 SPGSPGNTGLPFGPPGAGQGPGLKGERKETLQ 545
RESULT 9
CGHUTL
collagen alpha 1(III) chain precursor - human
N:Alternate names: procollagen alpha 1(III) chain
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1984 #sequence revision 01-Sep-1995 #text change 09-Jul-2004
C:Accession: S05272; S04642; PE0011; S01726; S04887; A90399; A94562; I51868; S59511; A90
R:Prockop, D.J.
submitted to the EMBL Data Library, February 1989
A:Reference number: S05272
A:Accession: S05272
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1240, 'V', 1242-1466 <PRC>
A:Cross-references: UNIPROT:P02461; UNIPARC:UPI0000000CDE; EMBL:X14420; NID:G30057; PIDN

R:Ala-Kokko, L.; Kontusaari, S.; Baldwin, C.T.; Kuivaniemi, H.; Prockop, D.J.
Biochem. J. 260, 509-516, 1989
A:Title: Structure of cDNA clones coding for the entire prepro-alpha(III) chain of hum
erences.
A:Reference number: S04642; MUID:89350838; PMID:2764886
A:Accession: S04642
A:Molecule type: mRNA
A:Residues: 1-1196 <ALA>
A:Cross-references: UNIPARC:UPI0000173B80; EMBL:X14420; NID:G30057; PIDN:CAA32583.1; PI
A:Note: the complete sequence is not shown
R:Benson-Chanda, V.; Su, M.W.; Weil, D.; Chu, M.L.; Ramirez, F.
Gene 78, 255-265, 1989
A:Title: Cloning and analysis of the 5' portion of the human type-III procollagen gene
A:Reference number: PE0011; MUID:89378752; PMID:2777083
A:Accession: PE0011
A:Molecule type: DNA
A:Residues: 1-176 <BEN>
A:Cross-references: UNIPARC:UPI000016A703; GB:M26939; NID:G180813; PIDN:AAAS2040.1; PID
R:Toman, P.D.; Ricca, G.A.; de Crombrughe, B.
Nucleic Acids Res. 16, 7201, 1988
A:Title: Nucleotide sequence of a cDNA coding for the amino-terminal region of human pr
A:Reference number: S01726; MUID:88303360; PMID:3405773
A:Accession: S01726
A:Molecule type: mRNA
A:Residues: 1-170 <TOM>
A:Cross-references: UNIPARC:UPI000016A706; EMBL:X07240; NID:G30060; PIDN:CAA30229.1; PI
A:Note: the authors translated the codon CAG for residue 154 as His
R:Janeczko, R.A.; Ramirez, F.
Nucleic Acids Res. 17, 6742, 1989
A:Title: Nucleotide and amino acid sequences of the entire human alpha-1 (III) collagen
A:Reference number: S04887; MUID:89386015; PMID:2780304
A:Accession: S04887
A:Molecule type: mRNA
A:Residues: 149-163, 'G', 164-240, 'D', 242-471, 'D', 473-487, 'L', 489, 'S', 491-613, 'Y', 615-634
A:Cross-references: UNIPARC:UPI000016A61C; EMBL:X15332; NID:G29545; PIDN:CAA33387.1; PI
A:Note: the authors' translation of residues 905-932 is inconsistent with the nucleotid
R:Seyzer, J.M.; Kang, A.H.
Biochemistry 16, 1158-1164, 1977
A:Title: Covalent structure of collagen: amino acid sequence of cyanogen bromide peptid
A:Reference number: A90399; MUID:77134724; PMID:557335
A:Accession: A90399
A:Molecule type: protein
A:Residues: 'V', 169-225, 229-232, 'P', 234-292, 'D', 294-398 <SEV1>
A:Cross-references: UNIPARC:UPI0000173B81
A:Experimental source: liver
A:Note: sequence corrected by A94562; attachment of 2-O-alpha-D-glucosyl-O-beta-D-galac
R:Seyzer, J.M.
submitted to the Atlas, December 1977
A:Reference number: A94562
A:Accession: A94562
A:Molecule type: protein
A:Residues: 'V', 169-225, 229-277, 'A', 279-292, 'D', 294, 'S', 296-398 <SEY2>
A:Cross-references: UNIPARC:UPI0000173B82
A:Experimental source: liver
A:Note: author submitted corrections to A90399
R:Milewicz, D.M.; Witz, A.M.; Smith, A.C.; Manchester, D.K.; Waldstein, G.; Byers, P.H.
Am. J. Hum. Genet. 53, 62-70, 1993
A:Title: Parental somatic and germ-line mosaicism for a multiexon deletion with unusual
spring.
A:Reference number: I51868; MUID:93304430; PMID:8317500
A:Accession: I51868
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 186-194 <MIL>
A:Cross-references: UNIPARC:UPI0000000B14; GB:S62925; NID:G386425; PIDN:AAD13937.1; PID
R:Chiodo, A.A.; Silience, D.O.; Cole, W.G.; Bateman, J.F.
Biochem. J. 311, 939-943, 1995
A:Title: Abnormal type III collagen produced by an exon-17-skipping mutation of the COL3
A:Reference number: S59511; MUID:96067614; PMID:7487954
A:Accession: S59511
A:Molecule type: mRNA
A:Residues: 302-423 <CHI>
A:Cross-references: UNIPARC:UPI0000173B83; GB:S79877; NID:G1195576; PIDN:AAB35615.1; PI

R;Seyer, J.M.; Kang, A.H.
Biochemistry 17, 3404-3411, 1978
A>Title: Covalent structure of collagen: amino acid sequence of five consecutive CNBr peptides
A/Reference number: A90414; MUID:79000343; PMID:687591
A/Accession: A90414
A/Molecule type: protein
A/Residues: 399-675, 'N', 677-727 <SEV3>
A/Cross-references: UNIPARC:UPI0000173884
A/Experimental source: liver
A/Title: Isolation of cDNA and genomic clones encoding human pro-alpha(I) collagen. P.
J. Biol. Chem. 266, 5256-5259, 1991
A/Reference number: I55349; MUID:91161621; PMID:1672129
A/Accession: I55349
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 537-605 <LEE>
A/Cross-references: UNIPARC:UPI0000004A2; GB:M59312; NID:g180815; PIDN:AAA52041.1; PID:
R;Seyer, J.M.; Mainardi, C.; Kang, A.H.
Biochemistry 19, 1583-1589, 1980
A>Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-CB5 from ty
A/Reference number: A90438; MUID:80198282; PMID:6246925
A/Accession: A90438
A/Molecule type: protein
A/Residues: 728-895, 'A', 897-964 <SEV4>
A/Cross-references: UNIPARC:UPI0000173885
A/Experimental source: liver
R;Cole, W.G.; Chiodo, A.A.; Lamanche, S.R.; Janeczko, R.; Ramirez, F.; Dahl, H.H.M.; Chan
J. Biol. Chem. 265, 17070-17077, 1990
A>Title: A base substitution at a splice site in the COL3A1 gene causes exon skipping an
A/Reference number: A93303; MUID:91009133; PMID:2145268
A/Accession: A93303
A/Molecule type: mRNA
A/Residues: 861-1015 <COL>
A/Cross-references: UNIPARC:UPI0000004A1; GB:J05617; GB:M55603; GB:M59227; NID:g180878;
A/Note: a mutant sequence with 942-977 spliced out from a patient with Ehlers-Danlos syn
R;Mankoo, B.S.; Dalgleish, R.
Nucleic Acids Res. 16, 2337, 1988
A>Title: Human pro alpha1(III) collagen: cDNA sequence for the 3' end.
A/Reference number: S02119; MUID:88189827; PMID:3357782
A/Accession: S02119
A/Status: translation not shown
A/Molecule type: mRNA
A/Residues: 950-1018, 'Y', 1020-1183, 'S', 1185-1466 <MAN>
A/Cross-references: UNIPARC:UPI0000173886; EMBL:X06700; NID:g30053; PIDN:CAA29886.1; PID
R;Seyer, J.M.; Kang, A.H.
Biochemistry 20, 2621-2627, 1981
A>Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-CB9 from ty
A/Reference number: A90446; MUID:81208139; PMID:7016180
A/Accession: A90446
A/Molecule type: protein
A/Residues: 965-979, 'A', 981-984, 'PS', 987, 'QN', 990-1096, 'P', 1098-1152, 'AT', 1155, 'S', 1157-
A/Cross-references: UNIPARC:UPI0000173887
A/Experimental source: liver
R;Loiol, H.R.; Brinker, J.M.; May, M.; Pihlajaniemi, T.; Morrow, S.; Rosenbloom, J.; Mye
Nucleic Acids Res. 12, 9383-9394, 1984
A>Title: Molecular cloning and carboxyl-propeptide analysis of human type III procollage
A/Reference number: A93551; MUID:85087944; PMID:6096827
A/Accession: A93551
A/Molecule type: mRNA
A/Residues: 1065-1155, 'P', 1157-1466 <LOI>
A/Cross-references: UNIPARC:UPI0000173888; EMBL:X01455; EMBL:X01742; NID:g29584; PIDN:CA
R;Mikulin, M.; Dalgleish, R.; Kluge-Beckerman, B.; Rennard, S.I.; Tolstoshev, P.; Brant
Biochemistry 25, 1408-1413, 1986
A>Title: Human type III collagen gene expression is coordinately modulated with the type
A/Reference number: I52393; MUID:86187804; PMID:3754462
A/Accession: I52393
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1161-1200 <MIS>
A/Cross-references: UNIPARC:UPI000016AGB5; GB:M13146; NID:g180415; PIDN:AAA52003.1; PID:
R;Emanuel, B.S.; Cannizzaro, L.A.; Seyer, J.M.; Myers, J.C.
Proc. Natl. Acad. Sci. U.S.A. 82, 3385-3389, 1985

A>Title: Human alpha 1(III) and alpha 2(V) procollagen genes are located on the long arm
A/Reference number: I59025; MUID:85216505; PMID:3858826
A/Accession: I79359
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1165-1196 <EMA>
A/Cross-references: UNIPARC:UPI000016AGB6; GB:M11134; NID:g180417; PIDN:AAA52004.1; PID:
R;Chu, M.L.; Weil, D.; de Wet, W.; Bernard, M.; Sippola, M.; Ramirez, F.
J. Biol. Chem. 260, 4357-4363, 1985
A>Title: Isolation of cDNA and genomic clones encoding human pro-alpha(III) collagen. P.
A/Reference number: A92516; MUID:85157600; PMID:2579949
A/Accession: A92516
A/Molecule type: DNA
A/Residues: 1176-1240, 'V', 1242-1356, 'P', 1358-1466 <CHU>
A/Cross-references: UNIPARC:UPI0000173889; GB:M10615; GB:M10793; GB:M10794; GB:M10795; G
A/Experimental source: liver
A/Note: the authors translated the codon TTC for residue 1057 as Tyr; the codons given f
C/Comment: Prolines and lysines at the third position of the tripeptide repeating unit (<
3-hydroxylated. About 15% of the lysines are 5-hydroxylated and some are subsequently O
C/Genetics:
A/Gene: GDB:COL3A1
A/Cross-references: GDB:I18729; OMIM:120180
A/Map position: 2q31-2q31
A/Introns: 27/1; 94/3; 111/3; 149/3; 176/3; 554/3; 587/3; 1175/3; 1275/1; 1337/3; 1418/3
A/Note: the list of introns is incomplete; defects in this gene can result in Ehlers-Dan
C/Complex: type III collagen is a homotrimer of monomers initially linked by disulfide b
er of their length, is formed with desmosine cross-links made from lysine and allysine r
C/Function:
A/Description: structural component of extracellular fibrous polymer that maintains inte
C/Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
C/Keywords: coiled coil; Ehlers-Danlos syndrome; extracellular matrix; glycoprotein; hyd
F/1-23/Domain: signal sequence #status predicted <SIG>
F/24-153/Domain: amino-terminal propeptide #status predicted <PRO>
F/31-91/Domain: von Willbrand factor type C repeat homology <WVC>
F/154-1221/Product: collagen alpha 1(III) chain #status predicted <MAT>
F/154-167/Region: amino-terminal nonhelical telopeptide
F/168-1196/Region: helical
F/1091-1093/Region: cell attachment (R-G-D) motif
F/1197-1221/Region: carboxyl-terminal nonhelical telopeptide
F/1222-1466/Domain: carboxyl-terminal propeptide #status predicted <CPR>
F/1238-1466/Domain: fibrillar collagen carboxyl-terminal homology <FCC>
F/24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F/153-154/Cleavage site: Pro-Gln (procollagen N-endopeptidase) #status predicted
F/154/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F/161,1212/Modified site: allysine (Lys) #status predicted
Query Match 9.5%; Score 124; DB 1; Length 1466;
Best Local Similarity 37.0%; Pred. No. 0, 0088;
Matches 40; Conservative 8; Mismatches 46; Indels 14; Gaps 5;
QY 22 LQLPAPSASIPKKGKQAQRQREVVDLY--NGWCLQGPGAGVFGDGSFGANGTPTGPG 79
Db 72 LDCPNP----EIPFGCCAVCPQPTAPTRPPNGQGPQKPGDGPFGIPGRNGDGPFG 127
QY 80 IPGRDGFKEGCECLRESF---ESWTPNYKQCSWSSLNLYGIDLKIA 124
Db 128 QPGSPGSPGPGIC--ESCPTGPQNYSPQYSDYDKS---GVAVGGLA 170
RESULT 10
S16366
collagen alpha 2(IV) chain precursor - pig roundworm
C/Species: Ascaris suum (pig roundworm)
C/Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C/Accession: S16366
R;Pettitt, J.; Kingston, I.B.
J. Biol. Chem. 266, 16149-16156, 1991
A>Title: The complete primary structure of a nematode alpha-2(IV) collagen and the parti
A/Reference number: S16366; MUID:91340768; PMID:1714907
A/Accession: S16366
A/Molecule type: mRNA
A/Residues: 1-1763 <JBI>

A;Cross-references: UNIPROT:P27393; UNIPARC:UPI0000126D40; GB:M67507; NID:G159648; PIDN:
C;Genetics:
A;Introns: 229/3; 266/3; 305/3; 360/3; 424/1; 489/1; 548/1; 656/3; 790/1; 891/1; 963/1;
C;Superfamily: collagen alpha 1(IV) chain
C;Keywords: alternative splicing; basement membrane; cell binding; coiled coil; disulfid
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-1763/Product: collagen alpha 2(IV) chain #status predicted <MAT>
F;27-42/Domain: non-collagenous NH1 #status predicted <NH1>
F;43-1529/Domain: collagenous #status predicted <COL>
F;197-199/Region: cell attachment (R-G-D) motif
F;1530-1763/Domain: carboxyl-terminal nonhelical, NC1 #status predicted <NC1>
F;1639-1763/Domain: repeat NC1 #status predicted <NC11>
F;31-34, 39, 41, 536, 539/disulfide bonds: interchain #status predicted
F;126/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;1593-1599, 1702-1709/disulfide bonds: #status predicted

Query Match 9.5%; Score 124; DB 2; Length 1763;
Best Local Similarity 25.8%; Pred. No. 0.011;
Matches 67; Conservative 24; Mismatches 93; Indels 76; Gaps 16;

QY 5 GPAASPORLRLGLLLQLLPAPSSASEIPKQKQAKLQREVVLDYNGM-CLQGPAGV 62
DB 1449 GPMGAF-GIRGEKGLPGLDGLPGSG-----PPGFAGAKGR-----DGFPGQPGMPGEKGA 1498

QY 63 PGRDGSPPGANGIPGTGIPGRDGFKEGKECLRESF---EESTWPNYKQCS-----W-- 111
DB 1499 PGLPGFPGLGEGIPGFLGPGSPGPPGPGPSYKDGFLLVKHSQTSVPPQCPGPMVKLWDG 1558

QY 112 -----SSLYNGIDGLKIAETFTTKMRSNSALRVLFSGSLRLKCRNAC-----CQR 156
DB 1559 YSLLYIEGNEKSHNQDLHGASC-----LSRFSTMPFLF-----CDVNVNVCNVRNDKSY 1609

QY 157 WYTFNGARCSGPLIEAIIYLDQSPENMSTINIHRTSSVEGLCGIGAGLVVDV---AI 213
DB 1610 WLST-----TAPIMMPV-----SEGGEIPYISRCACVCEA-PANVIAVHSQTI 1651

QY 214 WVTGCDSDYKGDASTGWSV 233
DB 1652 QIPNCFN-----GWNLSL 1663

RESULT 11
T15268
hypothetical protein F59E12.12 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T15268
R;Johnson, D.
submitted to the EMBL Data Library, May 1997
A;Description: The sequence of C. elegans cosmid F59E12.
A;Reference number: Z18318
A;Accession: T15268
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-311 <JOH>
A;Cross-references: UNIPROT:O01904; UNIPARC:UPI000007B0FF; EMBL:AF003386; NID:G2088833;
A;Experimental source: strain Bristol N2; clone F59E12
C;Genetics:
A;Gene: CESP:F59E12.12
A;Map position: 2
A;Introns: 24/2

Query Match 9.5%; Score 123.5; DB 2; Length 311;
Best Local Similarity 53.3%; Pred. No. 0.0017;
Matches 24; Conservative 4; Mismatches 12; Indels 5; Gaps 1;

QY 53 GMC-----LQGPAGVPRDGSPPGANGIPGTGIPGRDGFKEGKE 92
DB 227 GKCDENVNAQGGPPGPGFLPGDGLPTGTPGNFGQDGGQPAGE 271

RESULT 12

T26185
hypothetical protein W05B2.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T26185
R;Gardner, A.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z20166
A;Accession: T26185
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-304 <WIL>
A;Cross-references: UNIPROT:Q9U349; UNIPARC:UPI000007B910; EMBL:Z81138; PIDN:CAB03475.1
A;Experimental source: clone W05B2
C;Genetics:
A;Gene: CESP:W05B2.1
A;Map position: 3
A;Introns: 27/3

Query Match 9.4%; Score 123; DB 2; Length 304;
Best Local Similarity 33.0%; Pred. No. 0.0018;
Matches 30; Conservative 5; Mismatches 24; Indels 32; Gaps 2;

QY 3 PQGPAAASPORLRLGLLLQLLPAPSSASEIPKQKQAKLQREVVLDYNGMCLQGPAGV 62
DB 227 PQGP-----PGPAGSPGAPGPGQAG-----APGPKGP 254

QY 63 PGRDGSPPGANGIPGTGIPGRDGFKEGKEC 93
DB 255 SGAPQPGADGNPGAPGPGQSGGAGEKGIC 285

RESULT 13
T26184
hypothetical protein W05B2.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T26184
R;Gardner, A.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z20166
A;Accession: T26184
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-304 <WIL>
A;Cross-references: UNIPROT:Q9XVG3; UNIPARC:UPI0000079930; EMBL:Z81138; PIDN:CAB03474.1
A;Experimental source: clone W05B2
C;Genetics:
A;Gene: CESP:W05B2.6
A;Map position: 3
A;Introns: 27/3

Query Match 9.4%; Score 123; DB 2; Length 304;
Best Local Similarity 33.0%; Pred. No. 0.0018;
Matches 30; Conservative 5; Mismatches 24; Indels 32; Gaps 2;

QY 3 PQGPAAASPORLRLGLLLQLLPAPSSASEIPKQKQAKLQREVVLDYNGMCLQGPAGV 62
DB 227 PQGP-----PGPAGSPGAPGPGQAG-----APGPKGP 254

QY 63 PGRDGSPPGANGIPGTGIPGRDGFKEGKEC 93
DB 255 SGAPQPGADGNPGAPGPGQSGGAGEKGIC 285

RESULT 14
S22917
collagen alpha 5(IV) chain precursor, renal splice form - human
N;Alternate names: procollagen alpha 5(IV) chain
N;Contains: collagen alpha 5(IV) chain precursor, leukocyte splice form
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C;Accession: S22917; A54365; A57079; A37122; I54317; A34850; S18850; I56971; I76598; A3-

Db 1307 DQPPGLQGNPRGRLNGMKGDPLPGVPGFFPCMKG 1342

RESULT 15

T24482
 hypothetical protein T05A1.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T24482
 R:Illoyd, C.
 submitted to the EMBL Data Library, December 1995
 A:Reference number: Z19897
 A:Accession: T24482
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-300 <WIL>
 A:Cross-references: UNIPROT:Q22183; UNIPARC:UPI00000765DA; EMBL:Z68219; PIDN:CAA92476.1;
 A:Experimental source: clone T05A1
 C:Genetics:
 A:Gene: CESP:T05A1.2
 A:Map position: 4
 A:Introns: 7/3

Query Match 9.4%; Score 122.5; DB 2; Length 300;
 Best Local Similarity 31.5%; Pred. No. 0.002; Mismatches 28; Indels 37; Gaps 5;
 Matches 35; Conservative 11;
 QY 3 PQGPAASPOR--LRGILLILLLLQLPAPSSASIPKQKQAQLRQREVVDLYNGMCLOGP- 59
 Db 189 PHGPNCHPKPGSQG-----PPGPPGHSDPKPGQPGQFGRAGP-----RGPR 231
 QY 60 ---AGVPGRDGSPGA-----NGIPGTPGIPGRDGFKEGKEC 93
 Db 232 GVAGIKGKGAPGSPGQPGRGPGGPGQDGPAGPGPLFGSDGTPGKIGC 282

Search completed: April 20, 2006, 09:50:28
 Job time : 21.6941 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2006, 09:35:42 ; Search time 151.835 Seconds
(without alignments)
1129.140 Million cell updates/sec

Title: US-10-634-108-4
Perfect score: 1303
Sequence: 1 MRQPAPASPQLRGLLL.....GDASTGNSVSRILIELPK 243

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1303	100.0	243	1 CTRH1_HUMAN	Q96c98 homo sapien
2	1225	94.0	245	1 CTRH1_RAT	Q8cg08 rattus norv
3	1210	92.9	245	1 CTRH1_MOUSE	Q9d1d6 mus musculu
4	992	76.1	258	2 Q5HZD2_XENTR	Q5hzd2 xenopus tro
5	951	73.0	191	2 Q4SMQ6_TETNG	Q4smq6 tetraodon n
6	928	71.2	231	2 Q6AXL0_BRARE	Q6axl0 brachydanio
7	139	10.7	565	2 Q8K036_MOUSE	Q8k036 mus musculu
8	135	10.4	1208	2 Q4RX03_TETNG	Q4rx03 tetraodon n
9	133.5	10.2	652	2 Q5TAT5_HUMAN	Q5tat5 homo sapien
10	133.5	10.2	683	2 Q5TAT4_HUMAN	Q5tat4 homo sapien
11	133.5	10.2	717	2 Q9NQ52_HUMAN	Q9ng52 homo sapien
12	133.5	10.2	717	2 Q5TAT6_HUMAN	Q5tat6 homo sapien
13	133	10.2	289	2 Q6LBS4_CAEBR	Q6lbs4 caenorhabdi
14	131.5	10.1	246	1 C1QC_MOUSE	Q02105 mus musculu
15	131.5	10.1	1472	2 Q9OZA0_CHICK	Q9oza0 gallus gall
16	130.5	10.0	705	2 Q4SZU5_TETNG	Q4szu5 tetraodon n
17	129	9.9	287	2 Q8CFR0_MOUSE	Q8cfr0 mus musculu
18	129	9.9	1752	2 Q07265_STRPU	Q07265 strongyloe
19	128	9.8	289	2 Q18799_CAEBL	Q18799 caenorhabdi
20	128	9.8	1333	2 Q4SK58_TETNG	Q4sk58 tetraodon n
21	127.5	9.8	289	2 Q6QKP0_CAEBR	Q60qp0 caenorhabdi
22	127.5	9.8	358	2 Q6MEY7_PARUW	Q6mey7 parachlamyd
23	127.5	9.8	1069	2 Q6LAN8_HUMAN	Q6lan8 homo sapien
24	127.5	9.8	1461	2 Q76045_HUMAN	Q76045 homo sapien
25	127.5	9.8	1464	1 COLA1_HUMAN	P02452 homo sapien
26	127.5	9.8	1464	2 Q8N473_HUMAN	Q8n473 homo sapien
27	127.5	9.8	1467	2 Q59F64_HUMAN	Q59f64 homo sapien
28	127	9.7	291	2 Q9NAR3_CAEBL	Q9nar3 caenorhabdi
29	127	9.7	328	2 Q6LXP0_CAEBR	Q6lxp0 caenorhabdi
30	127	9.7	1151	2 Q4RMT1_TETNG	Q4rmt1 tetraodon n
31	126.5	9.7	540	2 Q86Y22_HUMAN	Q86y22 homo sapien

32	126.5	9.7	568	2	Q8CD80_MOUSE	Q8cd80 mus musculu
33	126.5	9.7	739	2	Q70575_MOUSE	Q70575 mus musculu
34	126.5	9.7	747	2	Q6NW57_BRARE	Q6nw57 brachydanio
35	126.5	9.7	751	2	Q9R1N9_MOUSE	Q9r1n9 mus musculu
36	126	9.7	327	2	Q01799_CAEBL	Q01799 caenorhabdi
37	126	9.7	371	2	Q4RUE0_TETNG	Q4rue0 tetraodon n
38	126	9.7	1447	2	Q6P4U1_BRARE	Q6p4u1 brachydanio
39	126	9.7	1447	2	Q6U1J5_BRARE	Q6u1j5 brachydanio
40	126	9.7	1669	2	Q9QZS0_MOUSE	Q9qzs0 mus musculu
41	125.5	9.6	418	2	Q4SDT0_TETNG	Q4sdt0 tetraodon n
42	125.5	9.6	887	2	Q53EJ9_BOMMO	Q53ej9 bombyx mori
43	125.5	9.6	1877	2	Q4RWT3_TETNG	Q4rwt3 tetraodon n
44	125	9.6	142	2	Q4RXU1_TETNG	Q4rxu1 tetraodon n
45	125	9.6	538	2	Q53QQ1_HUMAN	Q53qq1 homo sapien

ALIGNMENTS

RESULT 1

ID	CTRH1_HUMAN	STANDARD;	PRT;	243 AA.
AC	Q96CG8; Q6UW91; Q8IX63;			
DT	10-MAY-2005 (Rel. 47, Created)			
DT	10-MAY-2005 (Rel. 47, Last sequence update)			
DT	13-SEP-2005 (Rel. 48, Last annotation update)			
DE	Collagen triple helix repeat-containing protein 1 precursor (NMTC1 protein).			
GN	Name=CTRH1; ORFNames=UNQ762/PRO1550;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND TISSUE SPECIFICITY.			
RC	TISSUE=Aorta;			
RX	PubMed=15618538; DOI=10.1161/01.RES.0000154262.07264.12;			
RA	Pyagay P., Heroult M., Wang Q., Lehnert W., Beiden J., Liaw L., Friesel R.E., Lindner V.;			
RA	"Collagen triple helix repeat containing 1, a novel secreted protein in injured and diseased arteries, inhibits collagen expression and promotes cell migration.";			
RT	Circ. Res. 96:261-268(2005).			
RL	[2]			
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 2).			
RA	Sanuki N., Fujiki K., Kanai A., Tanaka Y., Iwata T.;			
RA	"Novel polypeptide found in human cornea cDNA library.";			
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).			
RX	MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;			
RA	Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J., Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P., Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S., Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J., Lewis L., Liao D., Mark M.R., Robbie B., Sanchez C., Schoenfeld J., Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A., Vandlen R.L., Watanabe C., Wiand D., Woods K., Xie M.-H., Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D., Wood W.I., Godowski P.J., Gray A.M.;			
RA	"The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment.";			
RL	Genome Res. 13:2265-2270(2003).			
RN	[4]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).			
RC	TISSUE=Kidney;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			

Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Scapleton M., Soares M.B., Bonaldo M.F., Casavani T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toehiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Vallaloo D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield J.S.N., Krzywinski M.J., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[5]

PROTEIN SEQUENCE OF 31-45 (ISOFORM 1).
PubMed=15340161; DOI=10.1110/ps.04682504;
Zhang Z., Henzel W.J.;
"Signal peptide prediction based on analysis of experimentally
verified cleavage sites";
Protein Sci. 13:2819-2824(2004).

-1- FUNCTION: May act as a negative regulator of collagen matrix
deposition (By similarity).

-1- SUBCELLULAR LOCATION: Secreted; extracellular (By similarity).

-1- ALTERNATIVE PRODUCTS:
Event-Alternative splicing; Named isoforms=2;
Names=1;
IsoId=O96CG8-1; Sequence=Displayed;
Names=2;
IsoId=O96CG8-2; Sequence=VSP_013622, VSP_013623;
Note=No experimental confirmation available;
TISSUE SPECIFICITY: Isoform 1 is expressed in calcified
atherosclerotic plaque and chondrocyte-like cells.

-1- PTM: N-glycosylated (By similarity).

-1- SIMILARITY: Contains 1 collagen-like domain.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

EMBL; AY136825; AAN15749.1; -; mRNA.
EMBL; AF395488; AA017919.1; -; mRNA.
EMBL; AY358914; AAQ89273.1; -; mRNA.
EMBL; BC014245; AAH14245.1; -; mRNA.
Ensembl; ENSG00000164932; Homo sapiens.
HGNC; HGNC:18831; CTRHC1.
InterPro; IPR008161; Clg_helix.
InterPro; IPR008160; Collagen.
Pfam; PF01391; Collagen; 1.
ProDom; PD000007; Clg_helix; 1.
Alternative splicing; Collagen; Direct protein sequencing;
Extracellular matrix; Glycoprotein; Signal; Structural protein.

FT	SIGNAL	1	30
FT	CHAIN	31	243
FT			Collagen triple helix repeat-containing protein 1.
FT	DOMAIN	57	90
FT	CARBOHYD	186	
FT	VARSPPLIC	1	50
FT			Collagen-like. N-linked (GlcNAc...) (Potential). MRPGQAPQRLGRLGLLGLLPPSSASEIPKGGKQKAAQ LRQREVVDL -> MWPPGRSITVKLRKREKTVSRKLEMGPSA FQGLICGK (in isoform 2). /FTId=VSP_013622. K -> IYML (in isoform 2). /FTId=VSP_013623. G -> V (in Ref. 3).
FT	VARSPPLIC	243	243
FT	CONFLICT	73	73
FT	SEQUENCE	243 AA;	26224 MW; A11PFBIC66867F9 CRC64;

Query Match 100.0%; Score 1303; DB 1; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.2e-107;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

KW Collagen; Extracellular matrix; Glycoprotein; Signal;
 FT Structural protein. 1 32 Potential.
 FT SIGNAL 33 245 Collagen triple helix repeat-containing
 FT CHAIN 33 245 protein 1.
 FT DOMAIN 59 92 Collagen-like.
 FT CARBOHYD 188 188 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 245 AA; 26424 MW; 2296FD6DCDBA21F2 CRC64;

Query Match 94.0%; Score 1225; DB 1; Length 245;
 Best Local Similarity 94.7%; Pred. No. 1.8e-100;
 Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MHPQGPAAAPQRLRG--LILLILLQLPAPSSASEIPKQKQALRQREVVDLYNGMCLQG 58
 DB 1 MHPQGPAAAPQRLRG--LILLILLQLPAPSSASEIPKQKQALRQREVVDLYNGMCLQG 60

QY 59 PAGVPRDGPANGIPGTPGIPGRDGFKEGECLEFESFESWTPNYKQCSWSLNYGI 118
 DB 61 PAGVPRDGPANGIPGTPGIPGRDGFKEGECLEFESFESWTPNYKQCSWSLNYGI 120

QY 119 DLGKIAECTFTTMRNSALRVLFSGSLRLKCRNACCQRYWTFNGAECGSLPIEAIYVL 178
 DB 121 DLGKIAECTFTTMRNSALRVLFSGSLRLKCRNACCQRYWTFNGAECGSLPIEAIYVL 180

QY 179 DOGSEPMNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPYPKGDASTGWSVSRIII 238
 DB 181 DOGSEPMNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPYPKGDASTGWSVSRIII 240

QY 239 BELPK 243
 DB 241 BELPK 245

RESULT 3
 CTRH1 MOUSE
 ID CTRH1 MOUSE STANDARD; PRT; 245 AA.
 AC Q9DD1D6;
 DT 10-MAY-2005 (Rel. 47, Created)
 DT 10-MAY-2005 (Rel. 47, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Collagen triple helix repeat-containing protein 1 precursor.
 GN Name=Cthrc1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Nunata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada M.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,

RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 CC -!- FUNCTION: May act as a negative regulator of collagen matrix
 deposition (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted; extracellular (By similarity).
 CC -!- PTM: N-glycosylated (By similarity).
 CC -!- SIMILARITY: Contains 1 collagen-like domain.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.
 CC -----
 DR EMBL; AK003674; BAB22930.1; -; mRNA.
 DR Ensembl; ENSMUSG0000054196; Mus musculus.
 DR MGI; MGI:1915838; Cthrc1.
 DR InterPro; IPR008161; C1g helix.
 DR InterPro; IPR008160; Collagen.
 DR Pfam; PF01391; Collagen; 1.
 DR ProDom; PD000007; C1g_helix; 1.
 DR Collagen; Extracellular matrix; Glycoprotein; Signal;
 KW Structural protein.
 FT SIGNAL 1 32 Potential.
 FT CHAIN 33 245 Collagen triple helix repeat-containing
 FT DOMAIN 59 92 Collagen-like.
 FT CARBOHYD 188 188 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 245 AA; 26460 MW; 14951B87D8181A0E CRC64;

Query Match 92.9%; Score 1210; DB 1; Length 245;
 Best Local Similarity 93.5%; Pred. No. 3.9e-99;
 Matches 229; Conservative 4; Mismatches 10; Indels 2; Gaps 1;

QY 1 MRPQGPAAAPQRLRG--LILLILLQLPAPSSASEIPKQKQALRQREVVDLYNGMCLQG 58
 DB 1 MHPQGPAAAPQRLRG--LILLILLQLPAPSSASEIPKQKQALRQREVVDLYNGMCLQG 60

QY 59 PAGVPRDGPANGIPGTPGIPGRDGFKEGECLEFESFESWTPNYKQCSWSLNYGI 118
 DB 61 PAGVPRDGPANGIPGTPGIPGRDGFKEGECLEFESFESWTPNYKQCSWSLNYGI 120

QY 119 DLGKIAECTFTTMRNSALRVLFSGSLRLKCRNACCQRYWTFNGAECGSLPIEAIYVL 178
 DB 121 DLGKIAECTFTTMRNSALRVLFSGSLRLKCRNACCQRYWTFNGAECGSLPIEAIYVL 180

QY 179 DOGSEPMNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPYPKGDASTGWSVSRIII 238
 DB 181 DOGSEPMNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPYPKGDASTGWSVSRIII 240

QY 239 BELPK 243
 DB 241 BELPK 245

RESULT 4
 QSHZD2 XENTR
 ID QSHZD2 XENTR PRELIMINARY; PRT; 258 AA.
 AC Q5HZD2;
 DT 10-MAY-2005 (TremblRel. 30, Created)
 DT 10-MAY-2005 (TremblRel. 30, Last sequence update)
 DT 10-MAY-2005 (TremblRel. 30, Last annotation update)
 DE LOC548356 protein (fragment).
 GN Name=LOC548356;
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

```
OC Xenopodinae: Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.W., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek A.M., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RA Klein S., Gerhard D.S.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC089073; AAB09073.1; -; mRNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008161; Clg_helix.
DR Pfam; PF01391; Collagen; 1.
DR ProDom; PD000007; Clg_helix; 1.
FT NON_TER 1
SQ SEQUENCE 258 AA; 28195 MW; F2E13334A1BA2CCA CRC64;

Query Match 76.1%; Score 992; DB 2; Length 258;
Best Local Similarity 79.6%; Pred. No. 9.4e-80;
Matches 183; Conservative 22; Mismatches 20; Indels 4; Gaps 3;

Qy 16 LLLLLLLQLPAPSSASEIPKQKQKQRLQRE--VVDLYNGMCLQPGAGVGRDGPAG 73
Db 31 LVLCSALLPL-HSQPENQVKQRS-LQKELDIERYNGMVCQVGMGPGRDGTGPGYNG 88

Qy 74 IPGTPGIPGRDGFKEGKECLRESPEESWTNPKQCSWSSLYNIGDLGKIAECTFTKMS 133
Db 89 IPGTPGIPGRDGAKEGKCEKMSLEESWTNFKQCAWSALNYGDLGKIAECTFTKMS 148

Qy 134 NSALRVFSSGLRLKCRNACCORWYFTFNGECSPLEIALIYLDQGSPEMNSTINHR 193
Db 149 HSAURVWFSSGLRLCKTACCORWYFTFNGECAGPLPIALIIYLDQGSPEFNSTINHR 208

Qy 194 TSSVEGLCEGIGAGLVDAIVAGTCSDPKGDASTGWSVSRRIIEELPK 243
Db 209 TSTVEGLCEGISAGLVDAIVAGTCSDPKGDASTGWSVSRRIIEELPK 258

RESULT 5
Q4SMQ6 TETNG
ID Q4SMQ6 TETNG PRELIMINARY; PRT; 191 AA.
AC Q4SMQ6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 8 SCAP14545, whole genome shotgun sequence.
DE (Fragment).
GN ORFName6=GSTENG00015647001;
```

```
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bimont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; CAAE01014545; CAF98076.1; -; Genomic_DNA.
DR InterPro; IPR008161; Clg_helix.
DR ProDom; PD000007; Collagen.
KW Collagen.
FT NON_TER 1
FT NON_TER 191
SQ SEQUENCE 191 AA; 20525 MW; CAB7BF4F582A3AA6 CRC64;

Query Match 73.0%; Score 951; DB 2; Length 191;
Best Local Similarity 85.9%; Pred. No. 2.9e-76;
Matches 164; Conservative 17; Mismatches 10; Indels 0; Gaps 0;

Qy 53 GMCLOGAGVPGRDGSPCANGIPCTPGIPGRDGFKEGKECLRESPEESWTNPKQCSWS 112
Db 1 GSCIQGGAGTIPGRDGNPGANGIPGTPGIPGRDGLKGEKGVCEIFEEFPWPNYKQCAWN 60

Qy 113 SLNYGIDLGKIAECTFTKMSNSALRVLFSSGLRLKCRNACCORWYFTFNGAECSPLEI 172
Db 61 SLNYGIDLGKVCDCFTKLRSDSLRLVLFSSGLRLKCKNACCORWYFTFNGAECTGPLV 120

Qy 173 EALIYLDQGSPEMNSTINIHRTSSVEGLCEGIGAGLVDAIVAGTCSDPKGDASTGWS 232
Db 121 ESIYLDQGSPELNSTINIHRTSSVEGMCEGVKAGLVDAIVAGTCSDPKGDASTGWS 180

Qy 233 VSRRIIEELPK 243
Db 181 VSRVIIEELPK 191

RESULT 6
Q6AXLO BRARE
ID Q6AXLO BRARE PRELIMINARY; PRT; 231 AA.
AC Q6AXLO;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Zgc:101075.
GN ORFName5=zgc:101075;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
```

NCBI_TaxID=10090;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Mix FVB/N;
RC TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;
RC MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carininci P., Prange C.,
RA Rana S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting R., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Kryzinski M.I., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
FT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Mix FVB/N;
RC TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;
RC Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJP databases.
RM EMBL; BC034164; AAH34164.1; -; mRNA.
DR MG1:1277203; Coll13a1.
DR GO; GO:0005911; C:intercellular junction; IDA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR Pfam; PF01391; Collagen; 5.
DR ProDom; PD000007; Clg_helix; 1.
KW Collagen.
SQ SEQUENCE 565 AA; 56726 MW; DED3FF99D670195F CRC64;
Query Match 10.7%; Score 139; DB 2; Length 565;
Best Local Similarity 34.8%; Pred. No. 0.0013;
Matches 39; Conservative 12; Mismatches 35; Indels 26; Gaps 6
Qy 27 PSSASEI-----PKGKQKQALRQREVVDLYNG-----MCLQGAGVGRDGSPCA 71
Db 393 PGAAGGQPSGPKG-ARKEPGKGMVD--YNGSINEALQEIETLALMGPLGPGQTGGPGP 450
Qy 72 NGIP-----GTPGIPGRDGFKGKGECLRSFEESWTPNYKQCSWSSLYNG 117
Db 451 PPTPGQRGEGTGLFPFGHDGDKGPRG---KPGDGNAGWSRTPRKGNWASRNEG 499
RESULT 8
Q4RX03 TETNG
ID Q4RX03 TETNG PRELIMINARY; PRT; 1208 AA.
AC Q4RX03;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 11 SCAF14979, whole genome shotgun sequence.
GN ORFNames=GSTENG00027621001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN NUCLEOTIDE SEQUENCE.
RP Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,

```
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lucifalla G., Dossat C., Segurens B.,
RA Dasilva C., Saïanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Bottolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Guzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet P., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Landier E.S., Weissbach J., Roest Crolius H.,
RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01014979; CAG07079.1; -; Genomic_DNA.
DR InterPro; IPR008161; Cig_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR000885; Fib_collagen_C.
DR Pfam; PF01410; COLFI; 1.
DR ProDom; PD000007; Cig_helix; 7.
DR ProDom; PD002078; Fib_collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
KW Collagen.
SQ SEQUENCE 1208 AA; 115322 MW; 1551D2793FE52F6A CRC64;

Query Match 10.4%; Score 135; DB 2; Length 1208;
Best Local Similarity 23.2%; Pred. No. 0.0071;
Matches 57; Conservative 24; Mismatches 63; Indels 102; Gaps 10;

QY 3 PQGPAAAPQR-----LRGLLLLLQLPAPSSASETPKQKQ----- 38
DB 958 PGPAGSGKDGKPGKIRGADGPPGQGDALRG-----PAGPSGKGDGEDGDPVG 1008
QY 39 -----KAQLRQREVVDL-----YNGMCLQGPAGVPGRDGSPGANGIPGTPGIPGR 83
DB 1009 PPGPSGPGLAGQRIIVLPQGRGRRGPPG--LPGPSGPKQAGPGTGGDRGPPGVP 1066
QY 84 DGFGEKGECLRESPEESWTPNYKQCSWSLNYGIDLKIA-----ECTFTKRSNSAL 137
DB 1067 PGLTGPAGELGRE-----FNYGDD--SLAPNTAAIQMTFLRLSLTEA- 1106
QY 138 RVLFSGSLRLKRNACCORWYFTFNAGCSGLPIEAIYLDQSPENWSTNIHRTSSV 197
DB 1107 ----SQNLTYHCKN-----SVAYMDASTGNLKKAVLLQGSNDV 1140
QY 198 EGLCEG 203
DB 1141 EIRAEG 1146

RESULT 9
QSTAT5 HUMAN
ID QSTAT5_HUMAN PRELIMINARY; PRT; 652 AA.
AC QSTAT5;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Collagen, type XIII, alpha 1 (Fragment).
GN Name=COL13A1; ORFNames=RP11-26212.1-002;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]

Query Match 10.2%; Score 133.5; DB 2; Length 683;
Best Local Similarity 43.4%; Pred. No. 0.0049;
Matches 33; Conservative 8; Mismatches 16; Indels 19; Gaps 4;

QY 34 PKGKQKQALRQREVVDLYNG-----MCLQGPAGVPGRDGSPGANGIP-----G 76
DB 409 PKG-SKGEPPKGEVND-YNGNINEALQEIRTLALMPPGLPQIGPPGAPGIPGQKGBIG 466
QY 77 TPGIPGRDGFKEKGE 92
DB 467 LPPGPPGHGDEKGPGRK 482

RESULT 10
QSTAT4 HUMAN
ID QSTAT4_HUMAN PRELIMINARY; PRT; 683 AA.
AC QSTAT4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Collagen, type XIII, alpha 1
GN Name=COL13A1; ORFNames=RP11-26212.1-004;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]

NUCLEOTIDE SEQUENCE.
RP Lawlor S.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL138925; CA115452.1; -; Genomic DNA.
DR EMBL; AC024601; CA115452.1; JOINED; Genomic DNA.
DR EMBL; AC025426; CA115452.1; JOINED; Genomic DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008161; Cig_helix.
DR InterPro; IPR008160; Collagen.
DR Pfam; PF01391; Collagen; 8.
DR ProDom; PD000007; Cig_helix; 3.
KW Collagen.
FT NON_TER
SQ SEQUENCE 652 AA; 63277 MW; BC950B456DBF34A6 CRC64;

Query Match 10.2%; Score 133.5; DB 2; Length 652;
Best Local Similarity 43.4%; Pred. No. 0.0047;
Matches 33; Conservative 8; Mismatches 16; Indels 19; Gaps 4;

QY 34 PKGKQKQALRQREVVDLYNG-----MCLQGPAGVPGRDGSPGANGIP-----G 76
DB 366 PKG-SKGEPPKGEVND-YNGNINEALQEIRTLALMPPGLPQIGPPGAPGIPGQKGBIG 423
QY 77 TPGIPGRDGFKEKGE 92
DB 424 LPPGPPGHGDEKGPGRK 439

RESULT 10
QSTAT4 HUMAN
ID QSTAT4_HUMAN PRELIMINARY; PRT; 683 AA.
AC QSTAT4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Collagen, type XIII, alpha 1
GN Name=COL13A1; ORFNames=RP11-26212.1-004;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]

NUCLEOTIDE SEQUENCE.
RP Lawlor S.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL138925; CA115452.1; -; Genomic DNA.
DR EMBL; AC024601; CA115452.1; JOINED; Genomic DNA.
DR EMBL; AC025426; CA115452.1; JOINED; Genomic DNA.
DR Ensembl; ENSG00000197467; Homo sapiens.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008161; Cig_helix.
DR InterPro; IPR008160; Collagen.
DR Pfam; PF01391; Collagen; 9.
DR ProDom; PD000007; Cig_helix; 2.
KW Collagen.
SQ SEQUENCE 683 AA; 66377 MW; 3BB550F0D03062C30 CRC64;

Query Match 10.2%; Score 133.5; DB 2; Length 683;
Best Local Similarity 43.4%; Pred. No. 0.0049;
Matches 33; Conservative 8; Mismatches 16; Indels 19; Gaps 4;

QY 34 PKGKQKQALRQREVVDLYNG-----MCLQGPAGVPGRDGSPGANGIP-----G 76
DB 409 PKG-SKGEPPKGEVND-YNGNINEALQEIRTLALMPPGLPQIGPPGAPGIPGQKGBIG 466
QY 77 TPGIPGRDGFKEKGE 92
DB 467 LPPGPPGHGDEKGPGRK 482
```

```

RESULT 11
Q9NQ52_HUMAN
ID Q9NQ52 HUMAN PRELIMINARY; PRT; 717 AA.
AC Q9NQ52;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Type XIII collagen.
GN Name=COL13A1; Synonyms=COLXIII1A1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Snellman A., Tu H., Vaeisaenen T., Kvist A.P., Huhtala P.,
RL Pihlajaniemi T.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ293624; CAC00688.1; -; mRNA.
DR Ensembl; ENSG00000197467; Homo sapiens.
DR HGNC; HGNC:2190; COL13A1.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005515; P:protein binding; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR Pfam; PF01391; Collagen; 9.
DR ProDom; PD000007; Clg_helix; 3.
DR KW
SQ SEQUENCE 717 AA; 69964 MW; A311B9C7D3E87577 CRC64;

Query Match 10.2%; Score 133.5; DB 2; Length 717;
Best Local Similarity 43.4%; Pred. No. 0.0052;
Matches 33; Conservative 8; Mismatches 16; Indels 19; Gaps 4;

QY 34 PKGKQKQAQLRQREVVDLYNG-----MCLQGAGVPGRGSGPGANGIP-----G 76
DB 431 PKG-SKEPGKGWVD-YNGNINEALQEIRTLALMGPPGLPQGIQPGGAPGIPQKGEIG 488
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 77 TPGIPGRDGFKGEGKE 92
DB 489 LPGPPGHGEGKGRGK 504
||| ||| ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
Q5TAT6_HUMAN
ID Q5TAT6 HUMAN PRELIMINARY; PRT; 717 AA.
AC Q5TAT6;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Collagen, type XIII, alpha 1.
GN Name=COL13A1; ORFNames=RP11-26212.1-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lawlor S.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL138925; CAIL15450.1; -; Genomic DNA.
DR EMBL; AC024601; CAIL15450.1; JOINED; Genomic DNA.
DR EMBL; AC025426; CAIL15450.1; JOINED; Genomic DNA.
DR HGNC; HGNC:2190; COL13A1.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.

```

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ICR; TISSUE=Macrophage;
RX MEDLINE=93011118; PubMed=1396691;
RA Petty F., Reid K.B.M., Loos M.;
RT "Isolation, sequence analysis and characterization of cDNA clones
RT coding for the C chain of mouse C1q. Sequence similarity of complement
RT subcomponent C1q, collagen type VII and type X and precerebellin.";
RT Eur. J. Biochem. 209:129-134(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BAUB/c; TISSUE=Liver;
RX MEDLINE=96186528; PubMed=8606057; DOI=10.1007/s002510050077;
RA Petty F., McClive P.J., Botto M., Morley B.J., Morahan G., Loos M.;
RT "The mouse C1q genes are clustered on chromosome 4 and show
RT conservation of gene organization.";
RL Immunogenetics 43:370-376(1996).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=FVB/N; TISSUE=Colon, and Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T.J., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Baha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.T., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: C1q associates with the proenzymes C1r and C1s to yield
CC C1, the first component of the serum complement system. The
CC collagen-like regions of C1q interact with the Ca(2+)-dependent
CC C1r(2)C1s(2) proenzyme complex, and efficient activation of C1
CC takes place on interaction of the globular heads of C1q with the
CC FC regions of IgG or IgM antibody present in immune complexes.
CC -!- SUBUNIT: C1 is a calcium-dependent trimolecular complex of C1q, R
CC and S in the molar ratio of 1:2:2. C1q subcomponent is composed
CC of nine subunits, six of which are disulfide-linked dimers of the
CC A and B chains, and three of which are disulfide-linked dimers of
CC the C chain.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 C1q domain.
CC -!- SIMILARITY: Contains 1 collagen-like domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; X66295; CAA46993.1; -; mRNA.
CC EMBL; X92960; CAA63535.1; -; Genomic_DNA.
CC EMBL; BC043945; AAH43945.1; -; mRNA.
CC EMBL; BC054443; AAH54443.1; -; mRNA.
CC PIR; S29328; S29328.
CC HSP; O60994; 1C28.
CC Ensembl; ENSMUSG00000036896; Mus musculus.
CC MGI; MGI:88225; C1qg.

DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001073; C1q_helix.
DR InterPro; IPR008161; C1q_helix.
DR Pfam; PF00386; C1q; 1.
DR PRINTS; PF01391; Collagen; 1.
DR PRODOM; PR000007; COMPLEMENTC1Q.
DR PROSITE; PS0871; C1q; 1.
KW Collagen; Complement pathway; Hydroxylation; Immune response;
KW Innate immunity; Plasma; Repeat; Signal.
FT SIGNAL 1 29
FT CHAIN 30 246 Complement C1q subcomponent, C chain.
FT DOMAIN 32 113 Collagen-like.
FT DOMAIN 116 246 C1q.
FT MOD_RES 37 37 Hydroxyproline (By similarity).
FT MOD_RES 40 40 Hydroxyproline (By similarity).
FT MOD_RES 43 43 Hydroxyproline (By similarity).
FT MOD_RES 61 61 Hydroxyproline (By similarity).
FT MOD_RES 64 64 Hydroxyproline (By similarity).
FT MOD_RES 73 73 5-hydroxylysine (By similarity).
FT MOD_RES 76 76 5-hydroxylysine (By similarity).
FT MOD_RES 79 79 Hydroxyproline (By similarity).
FT MOD_RES 82 82 Hydroxyproline (By similarity).
FT MOD_RES 97 97 Hydroxyproline (By similarity).
FT MOD_RES 100 100 Hydroxyproline (By similarity).
FT MOD_RES 106 106 Hydroxyproline (By similarity).
FT MOD_RES 109 109 Hydroxyproline (By similarity).
FT DISULFID 33 33 Interchain (with other C chain) (By similarity).
FT SQ SEQUENCE 246 AA; 25966 MW; 2F79EA1274BCB8E0 CRC64;
Query Match 10.1%; Score 131.5; DB 1; Length 246;
Best Local Similarity 38.5%; Pred. NO. 0.0022;
Matches 35; Conservative 5; Mismatches 32; Indels 19; Gaps 2;
QY 5 GPAAAPQRLGLLLLLQLLPAPSSASBIPKGKQAKLRQREVVDVLDNGMCLQGPAVGPG 64
DB 4 GPSQCPQGLCLLLFLALPERSQAS-----ACGYIPGMFGPGAPG 47
QY 65 RDGS---PGANGIPGTGIPGRDGFKGKGE 92
DB 48 KGDHGLQGKPGEGPAPVGTQGPKGQKGE 78
RESULT 15
Q90ZA0 CHICK
ID Q90ZA0 CHICK PRELIMINARY; PRT; 1472 AA.
AC Q90ZA0
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Collagen type XX alpha 1 precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21303548; PubMed=11274142; DOI=10.1074/jbc.M009912200;
RA Koch M., Foley J.E., Hahn R., Zhou P., Burgeson R.E., Gerecke D.R.,
RA Gordon M.K.;
RT "alpha 1(Xx) collagen, a new member of the collagen subfamily, fibril-
RT associated collagens with interrupted triple helices.";
RL J. Biol. Chem. 276:23120-23126(2001).
DR EMBL; AF312825; AAK58847.1; -; mRNA.
DR HSSP; Q96KE7; 1FNA.
DR Ensembl; ENSGALG00000005797; Gallus gallus.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.

DR GO; GO:0006817; P-phosphate transport; IEA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01391; Collagen; 4.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR00453; VWFADOMAIN.
DR ProDom; PD000007; Clg_helix; 4.
DR SMART; SM00060; FN3; 6.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00853; FN3; 6.
DR PROSITE; PS0234; VWF_A; 1.
KW Collagen; Signal.
FT SIGNAL 1 28 Potential.
SQ SEQUENCE 1472 AA; 156904 MW; 5361611579C56EFD CRC64;

Query Match 10.1%; Score 131.5; DB 2; Length 1472;
Best Local Similarity 34.5%; Pred. NO. 0.018;
Matches 38; Conservative 7; Mismatches 26; Indels 39; Gaps 6;

QY 3 PQGPAASPORLGLLLLLLLQLP-----APSSASEI-----PKGKQAQLRQREV 47
Db 1161 PFGPPGSPGR-RG-----PQGEQEPGPKGEPGKVGPGAPGPGSQGSGPSQ--- 1207

QY 48 VDLNMGMLQGPAGVP-----GRDGPANGIPGTPGIPGRDGFKEKG 91
Db 1208 -----GITIQGVPGPIKGEKGDYGSQMGQIPGVQGAFGRDGLQKAG 1252

Search completed: April 20, 2006, 09:49:21
Job time : 155.169 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2006, 09:49:46 ; Search time 27.6353 Seconds
(without alignments)
726.976 Million cell updates/sec

Title: US-10-634-108-4
Perfect score: 1303
Sequence: 1 MRPQGPAAAPQRLRGLLLLL.....GDASTGWNVSRIIEELPK 243

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1303	100.0	243	2	US-09-692-081-4
2	1300	99.8	243	2	US-09-489-847-205
3	1294	99.3	243	2	US-09-834-759-514
4	1294	99.3	243	2	US-10-012-231A-352
5	1294	99.3	243	2	US-10-012-231A-431
6	1294	99.3	243	2	US-10-015-389A-352
7	1294	99.3	243	2	US-10-015-389A-431
8	1294	99.3	243	2	US-10-006-768A-352
9	1294	99.3	243	2	US-10-006-768A-431
10	1294	99.3	243	2	US-10-015-671A-352
11	1294	99.3	243	2	US-10-015-671A-431
12	1294	99.3	243	2	US-10-015-393A-352
13	1294	99.3	243	2	US-10-015-393A-431
14	1294	99.3	243	2	US-10-011-833A-352
15	1294	99.3	243	2	US-10-011-833A-431
16	1294	99.3	243	2	US-10-006-041A-352
17	1294	99.3	243	2	US-10-006-041A-431
18	1294	99.3	243	2	US-10-012-064A-352
19	1294	99.3	243	2	US-10-012-064A-431
20	1294	99.3	243	2	US-10-076-622-514
21	1294	99.3	278	2	US-09-834-759-515
22	1294	99.3	278	2	US-10-076-622-515
23	1225	94.0	245	2	US-09-692-081-2
24	1225	94.0	277	2	US-09-692-081-5
25	1109.5	85.1	276	2	US-09-205-258-958
26	1109.5	85.1	276	2	US-10-004-860-958
27	1046	80.3	197	2	US-09-834-759-516

28	1046	80.3	197	2	US-10-076-622-516	Sequence 516, App
29	1046	80.3	232	2	US-09-834-759-517	Sequence 517, App
30	1046	80.3	232	2	US-10-076-622-517	Sequence 517, App
31	356	27.3	66	2	US-09-205-258-962	Sequence 962, App
32	356	27.3	66	2	US-10-004-860-962	Sequence 962, App
33	296	22.7	52	2	US-09-205-258-961	Sequence 961, App
34	296	22.7	52	2	US-10-004-860-961	Sequence 961, App
35	270	20.7	51	2	US-09-205-258-963	Sequence 963, App
36	270	20.7	51	2	US-10-004-860-963	Sequence 963, App
37	248	19.0	46	2	US-09-834-759-518	Sequence 518, App
38	248	19.0	46	2	US-10-076-622-518	Sequence 518, App
39	220.5	16.9	52	2	US-09-205-258-960	Sequence 960, App
40	220.5	16.9	52	2	US-10-004-860-960	Sequence 960, App
41	196	15.0	93	2	US-09-489-847-240	Sequence 240, App
42	196	15.0	93	2	US-09-489-847-362	Sequence 362, App
43	141	10.8	26	2	US-09-834-759-519	Sequence 519, App
44	141	10.8	26	2	US-10-076-622-519	Sequence 519, App
45	131.5	10.1	246	1	US-08-463-911-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-692-081-4
; Sequence 4, Application US/09692081
; Patent No. 6630325
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhard
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (2003036.0013)
; CURRENT APPLICATION NUMBER: US/09/692,081
; CURRENT FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-692-081-4

Query Match 100.0%; Score 1303; DB 2; Length 243;

Best Local Similarity 100.0%; Pred. No. 5.3e-125;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRPQGPAAAPQRLRGLLLLLLQLPAPSSASEIPKQKQALRQREVDLYNGMCLQSPA	60
DB	1	MRPQGPAAAPQRLRGLLLLLLQLPAPSSASEIPKQKQALRQREVDLYNGMCLQSPA	60
QY	61	GVPRGDSGANGIPGTGIPGRDGFKEGECLESFEESWTPNYKQCSWSLNYGIDL	120
DB	61	GVPRGDSGANGIPGTGIPGRDGFKEGECLESFEESWTPNYKQCSWSLNYGIDL	120
QY	121	GKIAECTTKWRSALRVLSGSLRLKRCRACCORWYFTFNGAECSGPLPIEAIYLDQ	180
DB	121	GKIAECTTKWRSALRVLSGSLRLKRCRACCORWYFTFNGAECSGPLPIEAIYLDQ	180
QY	181	GSPENWNTINHRITSSVEGLCEGIGAGLVDAIWIWGTCTDYPKGDASTGWNVSRIIEE	240
DB	181	GSPENWNTINHRITSSVEGLCEGIGAGLVDAIWIWGTCTDYPKGDASTGWNVSRIIEE	240
QY	241	LPK 243	
DB	241	LPK 243	

RESULT 2

US-09-489-847-205
; Sequence 205, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al

```
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031p1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 205
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-489-847-205

Query Match      99.8%; Score 1300; DB 2; Length 243;
Best Local Similarity 99.6%; Pred. No. 1.1e-124;
Matches 242; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MRPOGPAASPORLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
Db      1 MRPOGPAASPORLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60

Qy      61 GVPGRDGSFGANGIPGTPGIPGRDGFKEGKGECLRESFEESWTNPKQCSWSSLYNYGIDL 120
Db      61 GVPGRDGSFGANGIPGTPGIPGRDGFKEGKGECLRESFEESWTNPKQCSWSSLYNYGIDL 120

Qy      121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACQRYWFTFNGAECGSLPIEAIYLDQ 180
Db      121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACQRYWFTFNGAECGSLPIEAIYLDQ 180

Qy      181 GSPENSTINIHRTSSVEGLCEGIGAGLVDAIVWVGTCSDYPKGDASTGWNVSRIIEE 240
Db      181 GSPENSTINIHRTSSVEGLCEGIGAGLVDAIVWVGTCSDYPKGDASTGWNVSRIIEE 240

Qy      241 LPK 243
Db      241 LPK 243

RESULT 3
US-09-834-759-514
; Sequence 514, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 514
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-834-759-514

Query Match      99.8%; Score 1300; DB 2; Length 243;
Best Local Similarity 99.6%; Pred. No. 1.1e-124;
Matches 242; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MRPOGPAASPORLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
Db      1 MRPOGPAASPORLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60

Qy      61 GVPGRDGSFGANGIPGTPGIPGRDGFKEGKGECLRESFEESWTNPKQCSWSSLYNYGIDL 120
Db      61 GVPGRDGSFGANGIPGTPGIPGRDGFKEGKGECLRESFEESWTNPKQCSWSSLYNYGIDL 120

Qy      121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACQRYWFTFNGAECGSLPIEAIYLDQ 180
Db      121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACQRYWFTFNGAECGSLPIEAIYLDQ 180

Qy      181 GSPENSTINIHRTSSVEGLCEGIGAGLVDAIVWVGTCSDYPKGDASTGWNVSRIIEE 240
Db      181 GSPENSTINIHRTSSVEGLCEGIGAGLVDAIVWVGTCSDYPKGDASTGWNVSRIIEE 240

Qy      241 LPK 243
Db      241 LPK 243

RESULT 4
US-10-012-231A-352
; Sequence 352, Application US/10012231A
; Patent No. 6924355
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C23
; CURRENT APPLICATION NUMBER: US/10/012,231A
; CURRENT FILING DATE: 2002-06-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 352
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-012-231A-352

Query Match      99.3%; Score 1294; DB 2; Length 243;
Best Local Similarity 99.6%; Pred. No. 4.4e-124;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MRPOGPAASPORLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
Db      1 MRPOGPAASPORLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60

Qy      61 GVPGRDGSFGANGIPGTPGIPGRDGFKEGKGECLRESFEESWTNPKQCSWSSLYNYGIDL 120
Db      61 GVPGRDGSFGANGIPGTPGIPGRDGFKEGKGECLRESFEESWTNPKQCSWSSLYNYGIDL 120

Qy      121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACQRYWFTFNGAECGSLPIEAIYLDQ 180
Db      121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACQRYWFTFNGAECGSLPIEAIYLDQ 180
```

QY 181 GSPENSTINHTSSVEGLCEGAGLVDAIIVGTCTSDYKPGDASTGWNVSRIIEE 240
Db 181 GSPENSTINHTSSVEGLCEGAGLVDAIIVGTCTSDYKPGDASTGWNVSRIIEE 240
QY 241 LPK 243
Db 241 LPK 243

RESULT 5

US-10-012-231A-431
; Sequence 431, Application US/10012231A
; Patent No. 6924355
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan l.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC23
; CURRENT APPLICATION NUMBER: US/10/012,231A
; CURRENT FILING DATE: 2002-06-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 431
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-012-231A-431

Query Match 99.3%; Score 1294; DB 2; Length 243;
Best Local Similarity 99.6%; Pred. No. 4.4e-124;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRPGPAASPORLRLGLLLLLQLPAPSSASEIPKQKQAKQRLQREVVDLYNGMCLQGPA 60
Db 1 MRPGPAASPORLRLGLLLLLQLPAPSSASEIPKQKQAKQRLQREVVDLYNGMCLQGPA 60
QY 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKECLRESFEESWTPNYKQCSWSLNYGIDL 120
Db 61 GVPGRDGSFGANVIFGTGIPGRDGFKEGKECLRESFEESWTPNYKQCSWSLNYGIDL 120
QY 121 KIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYLDQ 180
Db 121 KIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYLDQ 180
QY 181 GSPENSTINHTSSVEGLCEGAGLVDAIIVGTCTSDYKPGDASTGWNVSRIIEE 240
Db 181 GSPENSTINHTSSVEGLCEGAGLVDAIIVGTCTSDYKPGDASTGWNVSRIIEE 240
QY 241 LPK 243
Db 241 LPK 243

RESULT 6

US-10-015-389A-352
; Sequence 352, Application US/10015389A
; Patent No. 6936436
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan l.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC48
; CURRENT APPLICATION NUMBER: US/10/015,389A
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 352
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-389A-352

Query Match 99.3%; Score 1294; DB 2; Length 243;
Best Local Similarity 99.6%; Pred. No. 4.4e-124;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRPGPAASPORLRLGLLLLLQLPAPSSASEIPKQKQAKQRLQREVVDLYNGMCLQGPA 60
Db 1 MRPGPAASPORLRLGLLLLLQLPAPSSASEIPKQKQAKQRLQREVVDLYNGMCLQGPA 60
QY 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKECLRESFEESWTPNYKQCSWSLNYGIDL 120
Db 61 GVPGRDGSFGANVIFGTGIPGRDGFKEGKECLRESFEESWTPNYKQCSWSLNYGIDL 120
QY 121 KIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYLDQ 180
Db 121 KIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYLDQ 180
QY 181 GSPENSTINHTSSVEGLCEGAGLVDAIIVGTCTSDYKPGDASTGWNVSRIIEE 240
Db 181 GSPENSTINHTSSVEGLCEGAGLVDAIIVGTCTSDYKPGDASTGWNVSRIIEE 240
QY 241 LPK 243
Db 241 LPK 243

RESULT 7

US-10-015-389A-431
; Sequence 431, Application US/10015389A
; Patent No. 6936436
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan l.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC48

```
; CURRENT APPLICATION NUMBER: US/10/015.389A
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 431
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-015-389A-431

Query Match          99.3%; Score 1294; DB 2; Length 243;
Best Local Similarity 99.6%; Pred. No. 4.4e-124;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MRQGPAAAPQRLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 60
Dy 1 MRQGPAAAPQRLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 60
Qy 61 GVPGRDGPANGIPGTPGIPGRDGFKEGKECLRESFEESWTNPKQCSWSSLYNYGIDL 120
Dy 61 GVPGRDGPANGIPGTPGIPGRDGFKEGKECLRESFEESWTNPKQCSWSSLYNYGIDL 120
Qy 121 GKIAECTFTKWRNSALRVLFGSLRLKCRNACQORWYFTFNGAECGSLPIEAIYLDQ 180
Dy 121 GKIAECTFTKWRNSALRVLFGSLRLKCRNACQORWYFTFNGAECGSLPIEAIYLDQ 180
Qy 181 GSPMNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWNVSRIIEE 240
Dy 181 GSPMNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWNVSRIIEE 240
Qy 241 LPK 243
Dy 241 LPK 243

RESULT 9
US-10-006-768A-431
; Sequence 431, Application US/10006768A
; Patent No. 6936697
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C10
; CURRENT APPLICATION NUMBER: US/10/006.768A
; CURRENT FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 477
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 431
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-006-768A-431

Query Match          99.3%; Score 1294; DB 2; Length 243;
Best Local Similarity 99.6%; Pred. No. 4.4e-124;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MRQGPAAAPQRLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 60
Dy 1 MRQGPAAAPQRLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 60
Qy 61 GVPGRDGPANGIPGTPGIPGRDGFKEGKECLRESFEESWTNPKQCSWSSLYNYGIDL 120
Dy 61 GVPGRDGPANGIPGTPGIPGRDGFKEGKECLRESFEESWTNPKQCSWSSLYNYGIDL 120
Qy 121 GKIAECTFTKWRNSALRVLFGSLRLKCRNACQORWYFTFNGAECGSLPIEAIYLDQ 180
Dy 121 GKIAECTFTKWRNSALRVLFGSLRLKCRNACQORWYFTFNGAECGSLPIEAIYLDQ 180
Qy 181 GSPMNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWNVSRIIEE 240
Dy 181 GSPMNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWNVSRIIEE 240
Qy 241 LPK 243
Dy 241 LPK 243

US-10-006-768A-352
; Sequence 352, Application US/10006768A
; Patent No. 6936697
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C10
; CURRENT APPLICATION NUMBER: US/10/006.768A
; CURRENT FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 477
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 352
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-006-768A-352

Query Match          99.3%; Score 1294; DB 2; Length 243;
Best Local Similarity 99.6%; Pred. No. 4.4e-124;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MRQGPAAAPQRLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 60
Dy 1 MRQGPAAAPQRLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 60
Qy 61 GVPGRDGPANGIPGTPGIPGRDGFKEGKECLRESFEESWTNPKQCSWSSLYNYGIDL 120
Dy 61 GVPGRDGPANGIPGTPGIPGRDGFKEGKECLRESFEESWTNPKQCSWSSLYNYGIDL 120
Qy 121 GKIAECTFTKWRNSALRVLFGSLRLKCRNACQORWYFTFNGAECGSLPIEAIYLDQ 180
Dy 121 GKIAECTFTKWRNSALRVLFGSLRLKCRNACQORWYFTFNGAECGSLPIEAIYLDQ 180
Qy 181 GSPMNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWNVSRIIEE 240
Dy 181 GSPMNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWNVSRIIEE 240
Qy 241 LPK 243
Dy 241 LPK 243
```

```
Db          241 LPK 243

RESULT 10
US-10-015-671A-352
; Sequence 352, Application US/10015671A
; Patent No. 6946263
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC47
; CURRENT APPLICATION NUMBER: US/10/015,671A
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 352
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-671A-352

Query Match          99.3%; Score 1294; DB 2; Length 243;
Best Local Similarity 99.6%; Pred. No. 4.4e-124;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 MRPQGPAAAPQRLRGLLLLLQLPAPSSASIPKKGKQAQLRQREVVDLYNGMCLQGPA 60
Db      1 MRPQGPAAAPQRLRGLLLLLQLPAPSSASIPKKGKQAQLRQREVVDLYNGMCLQGPA 60

QY      61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKGLRESFEESWTPNYKQCSWSLNYGIDL 120
Db      61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKGLRESFEESWTPNYKQCSWSLNYGIDL 120

QY      121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYLDQ 180
Db      121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYLDQ 180

QY      181 GSPENNSTINIHTSSVEGLCEGIGAGLVDAIWWGTCTSDYPKGDASTGWNVSRIIEE 240
Db      181 GSPENNSTINIHTSSVEGLCEGIGAGLVDAIWWGTCTSDYPKGDASTGWNVSRIIEE 240

QY      241 LPK 243
Db      241 LPK 243

RESULT 11
US-10-015-671A-431
; Sequence 431, Application US/10015671A
; Patent No. 6946263
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC46
; CURRENT APPLICATION NUMBER: US/10/015,393A
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 352
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-671A-431

Query Match          99.3%; Score 1294; DB 2; Length 243;
Best Local Similarity 99.6%; Pred. No. 4.4e-124;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 MRPQGPAAAPQRLRGLLLLLQLPAPSSASIPKKGKQAQLRQREVVDLYNGMCLQGPA 60
Db      1 MRPQGPAAAPQRLRGLLLLLQLPAPSSASIPKKGKQAQLRQREVVDLYNGMCLQGPA 60

QY      61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKGLRESFEESWTPNYKQCSWSLNYGIDL 120
Db      61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKGLRESFEESWTPNYKQCSWSLNYGIDL 120

QY      121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYLDQ 180
Db      121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYLDQ 180

QY      181 GSPENNSTINIHTSSVEGLCEGIGAGLVDAIWWGTCTSDYPKGDASTGWNVSRIIEE 240
Db      181 GSPENNSTINIHTSSVEGLCEGIGAGLVDAIWWGTCTSDYPKGDASTGWNVSRIIEE 240

QY      241 LPK 243
Db      241 LPK 243

RESULT 12
US-10-015-393A-352
; Sequence 352, Application US/10015393A
; Patent No. 6951737
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC46
; CURRENT APPLICATION NUMBER: US/10/015,393A
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 352
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-393A-431
```

```

; ORGANISM: Homo sapiens
US-10-015-393A-352

Query Match      99.3%; Score 1294; DB 2; Length 243;
Best Local Similarity 99.6%; Pred. No. 4.4e-124;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MRPGPAAAPORLRLGLLLLLQLPAPSSASEIPKGKQAKQLRQREVVDLYNGMCLQGPA 60
   |||||
Db 1 MRPGPAAAPORLRLGLLLLLQLPAPSSASEIPKGKQAKQLRQREVVDLYNGMCLQGPA 60
   |||||

Qy 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKECLRESFEESWTPTYKQCSWSSLYNGIDL 120
   |||||
Db 61 GVPGRDGSFGANVIPGTGIPGRDGFKEGKECLRESFEESWTPTYKQCSWSSLYNGIDL 120
   |||||

Qy 121 GKIAECTTKMRSNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYYLDQ 180
   |||||
Db 121 GKIAECTTKMRSNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYYLDQ 180
   |||||

Qy 181 GSPENNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWSVSRIIIEE 240
   |||||
Db 181 GSPENNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWSVSRIIIEE 240
   |||||

Qy 241 LPK 243
   |||
Db 241 LPK 243
   |||

RESULT 13
US-10-015-393A-431
; Sequence 431, Application US/10015393A
; Patent No. 6951737
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC26
; CURRENT APPLICATION NUMBER: US/10/015,393A
; CURRENT FILING DATE: 2002-06-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 431
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-015-393A-431

Query Match      99.3%; Score 1294; DB 2; Length 243;
Best Local Similarity 99.6%; Pred. No. 4.4e-124;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MRPGPAAAPORLRLGLLLLLQLPAPSSASEIPKGKQAKQLRQREVVDLYNGMCLQGPA 60
   |||||
Db 1 MRPGPAAAPORLRLGLLLLLQLPAPSSASEIPKGKQAKQLRQREVVDLYNGMCLQGPA 60
   |||||

Qy 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKECLRESFEESWTPTYKQCSWSSLYNGIDL 120
   |||||
Db 61 GVPGRDGSFGANVIPGTGIPGRDGFKEGKECLRESFEESWTPTYKQCSWSSLYNGIDL 120
   |||||

Qy 121 GKIAECTTKMRSNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYYLDQ 180
   |||||
Db 121 GKIAECTTKMRSNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYYLDQ 180
   |||||

Qy 241 LPK 243
   |||
Db 241 LPK 243
   |||

RESULT 14
US-10-011-833A-352
; Sequence 352, Application US/10011833A
; Patent No. 6951920
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC22
; CURRENT APPLICATION NUMBER: US/10/011,833A
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 352
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-011-833A-352

Query Match      99.3%; Score 1294; DB 2; Length 243;
Best Local Similarity 99.6%; Pred. No. 4.4e-124;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MRPGPAAAPORLRLGLLLLLQLPAPSSASEIPKGKQAKQLRQREVVDLYNGMCLQGPA 60
   |||||
Db 1 MRPGPAAAPORLRLGLLLLLQLPAPSSASEIPKGKQAKQLRQREVVDLYNGMCLQGPA 60
   |||||

Qy 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKECLRESFEESWTPTYKQCSWSSLYNGIDL 120
   |||||
Db 61 GVPGRDGSFGANVIPGTGIPGRDGFKEGKECLRESFEESWTPTYKQCSWSSLYNGIDL 120
   |||||

Qy 121 GKIAECTTKMRSNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYYLDQ 180
   |||||
Db 121 GKIAECTTKMRSNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYYLDQ 180
   |||||

Qy 181 GSPENNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWSVSRIIIEE 240
   |||||
Db 181 GSPENNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWSVSRIIIEE 240
   |||||

Qy 241 LPK 243
   |||
Db 241 LPK 243
   |||

RESULT 15
US-10-011-833A-431
; Sequence 431, Application US/10011833A
; Patent No. 6951920
```

Search completed: April 20, 2006, 09:52:00
Job time : 28.6353 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (C) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2006, 09:52:21 ; Search time 13.9765 Seconds
(without alignments)
765.051 Million cell updates/sec

Title: US-10-634-108-4

Perfect score: 1303

Sequence: 1 MRPQGPAAAPQLRGLLLLL.....GDASTGWSVRIIEELPK 243

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 225428 seqs, 44002918 residues

Total number of hits satisfying chosen parameters: 225428

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /SIDSS/ptodata/2/pubppa/US08_NEW_PUB.pep:*
- 2: /SIDSS/ptodata/2/pubppa/US06_NEW_PUB.pep:*
- 3: /SIDSS/ptodata/2/pubppa/US07_NEW_PUB.pep:*
- 4: /SIDSS/ptodata/2/pubppa/PCT_NEW_PUB.pep:*
- 5: /SIDSS/ptodata/2/pubppa/US09_NEW_PUB.pep:*
- 6: /SIDSS/ptodata/2/pubppa/US10_NEW_PUB.pep:*
- 7: /SIDSS/ptodata/2/pubppa/US11_NEW_PUB.pep:*
- 8: /SIDSS/ptodata/2/pubppa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1303	100.0	278	US-11-186-284-77	Sequence 77, Appl
2	1300	99.8	243	US-11-229-769-205	Sequence 205, App
3	1298	99.6	243	US-11-080-991-14	Sequence 14, Appl
4	1294	99.3	243	US-10-063-703-122	Sequence 122, App
5	1294	99.3	243	US-10-194-487-366	Sequence 366, App
6	1294	99.3	243	US-10-195-883-366	Sequence 366, App
7	1294	99.3	243	US-10-195-888-366	Sequence 366, App
8	1294	99.3	243	US-10-195-889-366	Sequence 366, App
9	1294	99.3	243	US-11-102-240-122	Sequence 122, App
10	1294	99.3	243	US-11-103-195-122	Sequence 122, App
11	1294	99.3	243	US-11-226-869-514	Sequence 514, App
12	1294	99.3	278	US-11-226-869-515	Sequence 515, App
13	1046	80.3	197	US-11-226-869-516	Sequence 516, App
14	1046	80.3	232	US-11-226-869-517	Sequence 517, App
15	248	19.0	46	US-11-226-869-518	Sequence 518, App
16	196	15.0	93	US-11-229-769-240	Sequence 240, App
17	196	15.0	93	US-11-229-769-362	Sequence 362, App
18	141	10.8	26	US-11-226-869-519	Sequence 519, App
19	129.5	9.9	585	US-10-508-440-2	Sequence 2, Appl
20	129	9.9	287	US-11-174-150-30	Sequence 30, Appl
21	127.5	9.8	1464	US-10-501-035-331	Sequence 331, App
22	127.5	9.8	1464	US-11-000-463-243	Sequence 243, App
23	127.5	9.8	1464	US-11-186-284-28	Sequence 28, Appl
24	127.5	9.8	1464	US-11-021-603-2	Sequence 2, Appl
25	127.5	9.8	1467	US-10-821-234-1096	Sequence 1096, Ap

ALIGNMENTS

RESULT 1

US-11-186-284-77
; Sequence 77, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MP01-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-186-284-77

Query Match 100.0%; Score 1303; DB 7; Length 278;
Best Local Similarity 100.0%; Pred. No. 2.7e-116;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPQGPAAAPQLRGLLLLLLQLPAPSSASEIPKGKQALRQREVVDLYNMCLOGPA 60
DB 36 MRPQGPAAAPQLRGLLLLLLQLPAPSSASEIPKGKQALRQREVVDLYNMCLOGPA 95
QY 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKECLRESFEESWTPNYKQCSWSLNYGIDL 120
DB 96 GVPGRDGSFGANGIPGTGIPGRDGFKEGKECLRESFEESWTPNYKQCSWSLNYGIDL 155
QY 121 GKIAECTTKMRNSALRVLFGSLRLKCRNACCORWYFTFNAGSCSGPLPIEAIYLDQ 180

Sequence 13, Appl
Sequence 12, Appl
Sequence 25, Appl
Sequence 30, Appl
Sequence 31, Appl
Sequence 2, Appl
Sequence 8, Appl
Sequence 964, App
Sequence 33, Appl
Sequence 375, App
Sequence 273, App
Sequence 376, App
Sequence 207, App
Sequence 280, App
Sequence 15, Appl
Sequence 4, Appl
Sequence 41, Appl
Sequence 1334, Ap
Sequence 1434, Ap
Sequence 911, App

26 126.5 9.7 540 6 US-10-508-440-13
27 125 9.6 551 7 US-11-096-070-12
28 124.5 9.6 141 7 US-11-243-689-25
29 124.5 9.6 288 7 US-11-135-855-30
30 124.5 9.6 303 7 US-11-135-855-31
31 124.5 9.6 303 7 US-11-258-647-2
32 124.5 9.6 1516 6 US-10-220-824-8
33 124 9.5 1166 6 US-10-821-234-964
34 124 9.5 1466 7 US-11-186-284-33
35 121 9.3 143 6 US-10-644-807-375
36 121 9.3 230 6 US-10-644-807-373
37 121 9.3 245 6 US-10-644-807-376
38 121 9.3 245 6 US-10-986-405-207
39 121 9.3 245 6 US-10-986-405-280
40 121 9.3 289 7 US-11-258-647-15
41 120 9.2 245 7 US-11-256-802-4
42 120 9.2 832 6 US-10-204-639-41
43 119 9.1 258 7 US-11-051-720-1334
44 119 9.1 258 7 US-11-051-720-1434
45 119 9.1 1767 6 US-10-995-561-911

Db 156 KIAECTTKMRSNSALRVLPSSGSLRLKCRNACCORWYFTFNGAECSGPLPIEAIYLDQ 215
Qy 181 GSPENNSTINHTSSVSEGLCEGAGLVDAIWIWGTCTSDYPKGDASTGWNVSRIIEE 240
Db 216 GSPENNSTINHTSSVSEGLCEGAGLVDAIWIWGTCTSDYPKGDASTGWNVSRIIEE 275
Qy 241 LPK 243
Db 276 LPK 278

RESULT 2

US-11-229-769-205
; Sequence 205, Application US/11229769
; Publication No. US20060079670A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PDI1
; CURRENT APPLICATION NUMBER: US/11/229,769
; CURRENT FILING DATE: 2005-09-20
; PRIOR APPLICATION NUMBER: 10/233,453
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 205
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-229-769-205

Query Match 99.8%; Score 1300; DB 7; Length 243;
Best Local Similarity 99.6%; Pred. No. 4.4e-116;
Matches 242; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRQGPAAAPQRLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
Db 1 MRQGPAAAPQRLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
Qy 61 GVPGRDGGPGANGIPGTGIPGRDGFKEGCECLRESFEESWTNPKQCSWSSLYNGIDL 120
Db 61 GVPGRDGGPGANGIPGTGIPGRDGFKEGCECLRESFEESWTNPKQCSWSSLYNGIDL 120
Qy 121 KIAECTTKMRSNSALRVLPSSGSLRLKCRNACCORWYFTFNGAECSGPLPIEAIYLDQ 180
Db 121 KIAECTTKMRSNSALRVLPSSGSLRLKCRNACCORWYFTFNGAECSGPLPIEAIYLDQ 180
Qy 181 GSPENNSTINHTSSVSEGLCEGAGLVDAIWIWGTCTSDYPKGDASTGWNVSRIIEE 240
Db 181 GSPENNSTINHTSSVSEGLCEGAGLVDAIWIWGTCTSDYPKGDASTGWNVSRIIEE 240
Qy 241 LPK 243
Db 241 LPK 243

RESULT 3

US-11-080-991-14

; Sequence 14, Application US/11080991
; Publication No. US20050266437A1
; GENERAL INFORMATION:
; APPLICANT: Velby, Petter Ole
; TITLE OF INVENTION: IDENTIFICATION, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/11/080,991
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: US/10/176,847
; PRIOR FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-080-991-14

Query Match 99.6%; Score 1298; DB 7; Length 243;
Best Local Similarity 99.6%; Pred. No. 6.9e-116;
Matches 242; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRQGPAAAPQRLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
Db 1 MRQGPAAAPQRLRGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
Qy 61 GVPGRDGGPGANGIPGTGIPGRDGFKEGCECLRESFEESWTNPKQCSWSSLYNGIDL 120
Db 61 GVPGRDGGPGANGIPGTGIPGRDGFKEGCECLRESFEESWTNPKQCSWSSLYNGIDL 120
Qy 121 KIAECTTKMRSNSALRVLPSSGSLRLKCRNACCORWYFTFNGAECSGPLPIEAIYLDQ 180
Db 121 KIAECTTKMRSNSALRVLPSSGSLRLKCRNACCORWYFTFNGAECSGPLPIEAIYLDQ 180
Qy 181 GSPENNSTINHTSSVSEGLCEGAGLVDAIWIWGTCTSDYPKGDASTGWNVSRIIEE 240
Db 181 GSPENNSTINHTSSVSEGLCEGAGLVDAIWIWGTCTSDYPKGDASTGWNVSRIIEE 240
Qy 241 LPK 243
Db 241 LPK 243

RESULT 4

US-10-063-703-122
; Sequence 122, Application US/10063703
; Publication No. US20060008901A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,703
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 122
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-703-122

Query Match 99.3%; Score 1294; DB 6; Length 243;

	Best Local Similarity	99.6%; Pred. No. 1.6e-115;	Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy	1	MRPQGPAASPORLRGLLLLLLQLLPAPSSASEIPKGKQAQLRQRREVVDLYNGMCLGPPA	60
Dd	1	MRPQGPAASPORLRGLLLLLLQLLPAPSSASEIPKGKQAQLRQRREVVDLYNGMCLGPPA	60
Qy	61	GVPGRDGS PGANGIPTPGIPIGRDGFKGEKECTRESPEESWTNPNYKCWSSSLNYGIDL	120
Dd	61	GVPGRDGS PGANVIPTPGIPIGRDGFKGEKECTRESPEESWTNPNYKCWSSSLNYGIDL	120
Qy	121	GKIACETFTKWRNSALRVLFSGSLURLKRNCACCORWYFTFNGAECSGPLPIEAIIVLDQ	180
Dd	121	GKIACETFTKWRNSALRVLFSGSLURLKRNCACCORWYFTFNGAECSGPLPIEAIIVLDQ	180
Qy	181	GSPENNSTINITHRTSSVEGLECGAGLVDAIVAVVGTCDSPKGDASTGNMSVSRIIEE	240
Dd	181	GSPENNSTINITHRTSSVEGLECGAGLVDAIVAVVGTCDSPKGDASTGNMSVSRIIEE	240
Qy	241	LPK 243 	
Dd	241	LPK 243 	

```

RESULT 5
US-10-194-487-366
; Sequence 366, Application US/10194487
; Publication No. US20060074226A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C312
; CURRENT APPLICATION NUMBER: US/10/194,487
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 366
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-194-487-366

```

US-10-194-487-366

```

Query Match          99.38; Score 1294; DB 6; Length 243;
Best Local Similarity 99.6%; Pred. No. 1.6e-115;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MRPGCPAASPORLRGLLLLLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
Db 1 MRPGCPAASPORLRGLLLLLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
Qy 61 GVPGRDGSPPANGIPGTPGIPGRDGFKEGKEGECLESPEBSWTNPNYKQCSWSSLNYGIDL 120
Db 61 GVPGRDGSPPANGVPGTPGIPGRDGFKEGKEGECLESPEBSWTNPNYKQCSWSSLNYGIDL 120
Qy 121 GKTAECTFTKMRNSALRVLFGSLSRLKCRNACCQRWFTFNGAECSGPLPIEATIIYLDQ 180
Db 121 GKTAECTFTKMRNSALRVLFGSLSRLKCRNACCQRWFTFNGAECSGPLPIEATIIYLDQ 180
Qy 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCDSDYPKGDASTGWNVSRIIEE 240
Db 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCDSDYPKGDASTGWNVSRIIEE 240
Qy 241 LPK 243
Db 241 LPK 243

RESULT 6
US-10-195-883-366
; Sequence 366, Application US/10195883
; Publication No. US20060073544A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1G323
; CURRENT APPLICATION NUMBER: US/10/195,883
; CURRENT FILING DATE: 2002-07-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 366
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-195-883-366

```

Query Match	99.3%;	Score 1294;	DB 6;	Length 243;
Best Local Similarity	99.6%;	Pred. No. 1.6e-115;		
Matches 242;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0
Qy	1	MRPGGPAASFORLGLLLLLLLLLLPAPSSASEIPKQKQALQRLQREVVDLYNGMCLQGPA	60	
Db	1	MRPGGPAASFORLGLLLLLLLLLQLPAPSSASEIPKQKQALQRLQREVVDLYNGMCLQGPA	60	
Qy	61	GVPGRDGSPGANGIPGTPGIPGRDGFKEGKECURESFEESWTNPNYKQCSWSSLYNGIDL	120	
Db	61	GVPGRDGSPGANVTPGTPGIPGRDGFKEGKECURESFEESWTNPNYKQCSWSSLYNGIDL	120	
Qy	121	GKTAECTFTKMRNSALRVLFSGSLRLKCRNACCQRYFTFNGAECGSPLEIAIYYLDQ	180	
Db	121	GKTAECTFTKMRNSALRVLFSGSLRLKCRNACCQRYFTFNGAECGSPLEIAIYYLDQ	180	
Qy	181	GSPEMNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPKGDASTGWNVSRIIEE	240	
Db	181	GSPEMNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPKGDASTGWNVSRIIEE	240	

```
QY      241 LPK 243
      |||
Db      241 LPK 243

RESULT 7
US-10-195-888-366
; Sequence 366, Application US/10195888
; Publication No. US20060073545A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C324
; CURRENT APPLICATION NUMBER: US/10/195,888
; CURRENT FILING DATE: 2002-07-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 366
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-195-888-366

Query Match      99.3%; Score 1294; DB 6; Length 243;
Best Local Similarity 99.6%; Pred. No. 1.6e-115;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 MRQGPAAAPQRLRGLLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
Db      1 MRQGPAAAPQRLRGLLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
QY      61 GVPGRDGGPGANGIPGTGIPGRDGFKEGKECLRESFEESWTNPKQCSWSSLYGIDL 120
Db      61 GVPGRDGGPGANGIPGTGIPGRDGFKEGKECLRESFEESWTNPKQCSWSSLYGIDL 120
QY      121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPTEAIYLDQ 180
Db      121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPTEAIYLDQ 180
QY      181 GSPENSTINHTSSVEGLCEGAGLDVDAIWWGTCSDPKGDASTGMNSVSRILIEE 240
Db      181 GSPENSTINHTSSVEGLCEGAGLDVDAIWWGTCSDPKGDASTGMNSVSRILIEE 240
QY      241 LPK 243
      |||
Db      241 LPK 243

RESULT 8
US-10-195-889-366
; Sequence 366, Application US/10195889
; Publication No. US20060074227A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C324
; CURRENT APPLICATION NUMBER: US/10/195,888
; CURRENT FILING DATE: 2002-07-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 366
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-195-888-366

Query Match      99.3%; Score 1294; DB 6; Length 243;
Best Local Similarity 99.6%; Pred. No. 1.6e-115;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 MRQGPAAAPQRLRGLLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
Db      1 MRQGPAAAPQRLRGLLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
QY      61 GVPGRDGGPGANGIPGTGIPGRDGFKEGKECLRESFEESWTNPKQCSWSSLYGIDL 120
Db      61 GVPGRDGGPGANGIPGTGIPGRDGFKEGKECLRESFEESWTNPKQCSWSSLYGIDL 120
QY      121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPTEAIYLDQ 180
Db      121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPTEAIYLDQ 180
QY      181 GSPENSTINHTSSVEGLCEGAGLDVDAIWWGTCSDPKGDASTGMNSVSRILIEE 240
Db      181 GSPENSTINHTSSVEGLCEGAGLDVDAIWWGTCSDPKGDASTGMNSVSRILIEE 240
QY      241 LPK 243
      |||
Db      241 LPK 243

RESULT 9
US-11-102-240-122
; Sequence 122, Application US/11102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
; FILE REFERENCE: P3230R1C106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; CURRENT FILING DATE: 2005-04-08
; Prior Application Number: 10/063662
; Prior Filing Date: 2002-05-07
; Prior Application Number: 10/068867
; Prior Filing Date: 2001-12-06
; Prior Application Number: PCT/US00/23328
; Prior Filing Date: 2000-08-24
; Prior Application Number: 60/170262
; Prior Filing Date: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 122
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-102-240-122

Query Match      99.3%; Score 1294; DB 7; Length 243;
Best Local Similarity 99.6%; Pred. No. 1.6e-115;
```

```
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MRPOGPAASPORLGLLLLLQLPAPSSASBIPKQKQAKQALRQREVVDLYNGMCLQGPA 60
DB 1 MRPOGPAASPORLGLLLLLQLPAPSSASBIPKQKQAKQALRQREVVDLYNGMCLQGPA 60
QY 61 GVPGRDGSFGANGIPCTPGIPGRDGFKGKGECLRESFEESWTPNPKQCSWSLSNYGIDL 120
DB 61 GVPGRDGSFGANGIPCTPGIPGRDGFKGKGECLRESFEESWTPNPKQCSWSLSNYGIDL 120
QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCRQRYFTFNGAECGSLPIEAIYLDQ 180
DB 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCRQRYFTFNGAECGSLPIEAIYLDQ 180
QY 181 GSPENNSTINIHTSSVEGLCEGAGLVDAIWDVVAIWGTCSDYKPGDASTGWSVSRIIIE 240
DB 181 GSPENNSTINIHTSSVEGLCEGAGLVDAIWDVVAIWGTCSDYKPGDASTGWSVSRIIIE 240
QY 241 LPK 243
DB 241 LPK 243
RESULT 10
US-11-103-195-122
; Sequence 122, Application US/11103195
; Publication No. US20060051779A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230C166C1
; CURRENT APPLICATION NUMBER: US/11/103,195
; CURRENT FILING DATE: 2005-04-11
; PRIOR APPLICATION NUMBER: 10/063743
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 122
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-103-195-122
Query Match 99.3%; Score 1294; DB 7; Length 243;
Best Local Similarity 99.6%; Pred. No. 1.6e-115;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MRPOGPAASPORLGLLLLLQLPAPSSASBIPKQKQAKQALRQREVVDLYNGMCLQGPA 60
DB 1 MRPOGPAASPORLGLLLLLQLPAPSSASBIPKQKQAKQALRQREVVDLYNGMCLQGPA 60
QY 61 GVPGRDGSFGANGIPCTPGIPGRDGFKGKGECLRESFEESWTPNPKQCSWSLSNYGIDL 120
DB 61 GVPGRDGSFGANGIPCTPGIPGRDGFKGKGECLRESFEESWTPNPKQCSWSLSNYGIDL 120
QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCRQRYFTFNGAECGSLPIEAIYLDQ 180
DB 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCRQRYFTFNGAECGSLPIEAIYLDQ 180
QY 181 GSPENNSTINIHTSSVEGLCEGAGLVDAIWDVVAIWGTCSDYKPGDASTGWSVSRIIIE 240
DB 181 GSPENNSTINIHTSSVEGLCEGAGLVDAIWDVVAIWGTCSDYKPGDASTGWSVSRIIIE 240
QY 241 LPK 243
DB 241 LPK 243
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MRPOGPAASPORLGLLLLLQLPAPSSASBIPKQKQAKQALRQREVVDLYNGMCLQGPA 60
DB 1 MRPOGPAASPORLGLLLLLQLPAPSSASBIPKQKQAKQALRQREVVDLYNGMCLQGPA 60
QY 61 GVPGRDGSFGANGIPCTPGIPGRDGFKGKGECLRESFEESWTPNPKQCSWSLSNYGIDL 120
DB 61 GVPGRDGSFGANGIPCTPGIPGRDGFKGKGECLRESFEESWTPNPKQCSWSLSNYGIDL 120
QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCRQRYFTFNGAECGSLPIEAIYLDQ 180
DB 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCRQRYFTFNGAECGSLPIEAIYLDQ 180
QY 181 GSPENNSTINIHTSSVEGLCEGAGLVDAIWDVVAIWGTCSDYKPGDASTGWSVSRIIIE 240
DB 181 GSPENNSTINIHTSSVEGLCEGAGLVDAIWDVVAIWGTCSDYKPGDASTGWSVSRIIIE 240
QY 241 LPK 243
DB 241 LPK 243
```

```
Db 181 GSPENNSTINIHTSSVEGLCEGAGLVDAIWDVVAIWGTCSDYKPGDASTGWSVSRIIIE 240
QY 241 LPK 243
Db 241 LPK 243
RESULT 11
US-11-226-869-514
; Sequence 514, Application US/11226869
; Publication No. US20060069054A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C14
; CURRENT APPLICATION NUMBER: US/11/226,869
; CURRENT FILING DATE: 2005-09-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 514
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-226-869-514
Query Match 99.3%; Score 1294; DB 7; Length 243;
Best Local Similarity 99.6%; Pred. No. 1.6e-115;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MRPOGPAASPORLGLLLLLQLPAPSSASBIPKQKQAKQALRQREVVDLYNGMCLQGPA 60
DB 1 MRPOGPAASPORLGLLLLLQLPAPSSASBIPKQKQAKQALRQREVVDLYNGMCLQGPA 60
QY 61 GVPGRDGSFGANGIPCTPGIPGRDGFKGKGECLRESFEESWTPNPKQCSWSLSNYGIDL 120
DB 61 GVPGRDGSFGANGIPCTPGIPGRDGFKGKGECLRESFEESWTPNPKQCSWSLSNYGIDL 120
QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCRQRYFTFNGAECGSLPIEAIYLDQ 180
DB 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCRQRYFTFNGAECGSLPIEAIYLDQ 180
QY 181 GSPENNSTINIHTSSVEGLCEGAGLVDAIWDVVAIWGTCSDYKPGDASTGWSVSRIIIE 240
DB 181 GSPENNSTINIHTSSVEGLCEGAGLVDAIWDVVAIWGTCSDYKPGDASTGWSVSRIIIE 240
QY 241 LPK 243
DB 241 LPK 243
RESULT 12
US-11-226-869-515
; Sequence 515, Application US/11226869
; Publication No. US20060069054A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C14
; CURRENT APPLICATION NUMBER: US/11/226,869
; CURRENT FILING DATE: 2005-09-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 515
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-226-869-515
```

Query Match 99.3%; Score 1294; DB 7; Length 278;
Best Local Similarity 99.6%; Pred. No. 1.9e-115;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRPGPAASPORLRLGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 60
DB 36 MRPGPAASPORLRLGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 95

QY 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGECLRESFEESWTNPKQCSWSSLYNGIDL 120
DB 96 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGECLRESFEESWTNPKQCSWSSLYNGIDL 155

QY 121 GKIAECTTKMRSNSALRVLSFGSLRLKCRNACCORWYFTNGAECSPLEIAIYLDQ 180
DB 156 GKIAECTTKMRSNSALRVLSFGSLRLKCRNACCORWYFTNGAECSPLEIAIYLDQ 215

QY 181 GSPENNSTINIHRTSSV 197
DB 276 LPK 278

RESULT 13

US-11-226-869-516
; Sequence 516, Application US/11226869
; Publication No. US20060069054A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C14
; CURRENT APPLICATION NUMBER: US/11/226,869
; CURRENT FILING DATE: 2005-09-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 516
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-226-869-516

Query Match 80.3%; Score 1046; DB 7; Length 197;
Best Local Similarity 99.5%; Pred. No. 4.7e-92;
Matches 196; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRPGPAASPORLRLGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 60
DB 1 MRPGPAASPORLRLGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 60

QY 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGECLRESFEESWTNPKQCSWSSLYNGIDL 120
DB 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGECLRESFEESWTNPKQCSWSSLYNGIDL 120

QY 121 GKIAECTTKMRSNSALRVLSFGSLRLKCRNACCORWYFTNGAECSPLEIAIYLDQ 180
DB 121 GKIAECTTKMRSNSALRVLSFGSLRLKCRNACCORWYFTNGAECSPLEIAIYLDQ 180

QY 181 GSPENNSTINIHRTSSV 197
DB 181 GSPENNSTINIHRTSSV 197

RESULT 14

US-11-226-869-517
; Sequence 517, Application US/11226869
; Publication No. US20060069054A1
; GENERAL INFORMATION:

; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C14
; CURRENT APPLICATION NUMBER: US/11/226,869
; CURRENT FILING DATE: 2005-09-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 517
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-226-869-517

Query Match 80.3%; Score 1046; DB 7; Length 232;
Best Local Similarity 99.5%; Pred. No. 5.7e-92;
Matches 196; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRPGPAASPORLRLGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 60
DB 36 MRPGPAASPORLRLGLLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQCPA 95

QY 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGECLRESFEESWTNPKQCSWSSLYNGIDL 120
DB 96 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGECLRESFEESWTNPKQCSWSSLYNGIDL 155

QY 121 GKIAECTTKMRSNSALRVLSFGSLRLKCRNACCORWYFTNGAECSPLEIAIYLDQ 180
DB 156 GKIAECTTKMRSNSALRVLSFGSLRLKCRNACCORWYFTNGAECSPLEIAIYLDQ 215

QY 181 GSPENNSTINIHRTSSV 197
DB 216 GSPENNSTINIHRTSSV 232

RESULT 15

US-11-226-869-518
; Sequence 518, Application US/11226869
; Publication No. US20060069054A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C14
; CURRENT APPLICATION NUMBER: US/11/226,869
; CURRENT FILING DATE: 2005-09-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 518
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-226-869-518

Query Match 19.0%; Score 248; DB 7; Length 46;
Best Local Similarity 100.0%; Pred. No. 5.9e-17;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 ECLCEGAGLVDAIWIWGTCTSDYKGDASTGWSVSRILIEELPK 243
DB 1 ECLCEGAGLVDAIWIWGTCTSDYKGDASTGWSVSRILIEELPK 46

Search completed: April 20, 2006, 09:57:47
Job time : 15.9765 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2006, 09:50:45 ; Search time 92.1176 Seconds
(without alignments)
1102.205 Million cell updates/sec

Title: US-10-634-108-4
Perfect score: 1303
Sequence: 1 MRQGPAAAPQRLRGLLLLL.....GDASTGWSVSRILIEELPK 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgm2_6/ptodata/1/pubpa/US07_PUBCOMB.pcp.*
2: /cgm2_6/ptodata/1/pubpa/US08_PUBCOMB.pcp.*
3: /cgm2_6/ptodata/1/pubpa/US09_PUBCOMB.pcp.*
4: /cgm2_6/ptodata/1/pubpa/US10A_PUBCOMB.pcp.*
5: /cgm2_6/ptodata/1/pubpa/US10B_PUBCOMB.pcp.*
6: /cgm2_6/ptodata/1/pubpa/US11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1303	100.0	243	3	US-09-829-472A-19
2	1303	100.0	243	4	US-10-045-992-4
3	1303	100.0	243	4	US-10-097-340-135
4	1303	100.0	243	4	US-10-295-027-50
5	1303	100.0	243	4	US-10-295-027-1173
6	1303	100.0	243	4	US-10-173-999-74
7	1303	100.0	243	4	US-10-058-270A-38
8	1303	100.0	243	4	US-10-634-108-4
9	1303	100.0	243	4	US-10-188-832-175
10	1303	100.0	243	5	US-10-939-233-4
11	1303	100.0	243	6	US-11-050-926-135
12	1303	100.0	278	4	US-10-097-340-131
13	1303	100.0	278	4	US-10-177-293-496
14	1303	100.0	278	4	US-10-301-822-77
15	1303	100.0	278	4	US-10-296-115-1261
16	1303	100.0	278	5	US-10-961-139-2
17	1303	100.0	278	6	US-11-050-926-131
18	1300	99.8	243	4	US-10-351-334-205
19	1300	99.8	278	4	US-10-060-036-4551
20	1298	99.6	243	4	US-10-176-847-14
21	1294	99.3	243	3	US-09-834-759-514
22	1294	99.3	243	3	US-09-938-418-7
23	1294	99.3	243	3	US-09-946-374-352
24	1294	99.3	243	3	US-09-946-374-431
25	1294	99.3	243	4	US-10-006-867-122
26	1294	99.3	243	4	US-10-052-586-366
27	1294	99.3	243	4	US-10-007-805-514

28	1294	99.3	243	4	US-10-063-547-122	Sequence 122, App
29	1294	99.3	243	4	US-10-063-551-122	Sequence 122, App
30	1294	99.3	243	4	US-10-174-590-366	Sequence 366, App
31	1294	99.3	243	4	US-10-176-758-366	Sequence 366, App
32	1294	99.3	243	4	US-10-175-737-366	Sequence 366, App
33	1294	99.3	243	4	US-10-063-616-122	Sequence 366, App
34	1294	99.3	243	4	US-10-174-581-366	Sequence 366, App
35	1294	99.3	243	4	US-10-176-483-366	Sequence 366, App
36	1294	99.3	243	4	US-10-176-749-366	Sequence 366, App
37	1294	99.3	243	4	US-10-176-914-366	Sequence 366, App
38	1294	99.3	243	4	US-10-176-915-366	Sequence 366, App
39	1294	99.3	243	4	US-10-063-569-122	Sequence 122, App
40	1294	99.3	243	4	US-10-063-513-122	Sequence 122, App
41	1294	99.3	243	4	US-10-063-515-122	Sequence 122, App
42	1294	99.3	243	4	US-10-063-512-122	Sequence 122, App
43	1294	99.3	243	4	US-10-173-706-366	Sequence 366, App
44	1294	99.3	243	4	US-10-175-738-366	Sequence 366, App
45	1294	99.3	243	4	US-10-175-752-366	Sequence 366, App

ALIGNMENTS

RESULT 1
US-09-829-472A-19
; Sequence 19, Application US/09829472A
; Publication No. US2004014862A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Breast Cancer
; FILE REFERENCE: 018501-001200US
; CURRENT APPLICATION NUMBER: US/09/829,472A
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 09/525,361
; PRIOR FILING DATE: 2000-03-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Ratent in Ver. 2.1
; SEQ ID NO 19
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: BCN4, ESTs, secreted protein
US-09-829-472A-19

Query Match	100.0%	Score 1303;	DB 3;	Length 243;
Best Local Similarity	100.0%	Pred. No. 1.6e-115;	Mismatches 0;	Indels 0; Gaps 0;
Matches 243;	Conservative 0;			
Qy	1	MRQGPAAAPQRLRGLLLLLLQLPAPSSASEIPKQKQKQALRQREVVDLYNGMCLQGP 60		
Db	1	MRQGPAAAPQRLRGLLLLLLQLPAPSSASEIPKQKQKQALRQREVVDLYNGMCLQGP 60		
Qy	61	GVPRDGSFGANGIPGTPGIPGRDGFKEGKECLRESEESWTPTNYKQCSWSLNYGIDL 120		
Db	61	GVPRDGSFGANGIPGTPGIPGRDGFKEGKECLRESEESWTPTNYKQCSWSLNYGIDL 120		
Qy	121	GKIAECTTKRNSALRVLPFGSISRLKCRNACORWYFTFGACSCGPLPTEAIIYLDQ 180		
Db	121	GKIAECTTKRNSALRVLPFGSISRLKCRNACORWYFTFGACSCGPLPTEAIIYLDQ 180		
Qy	181	GSPENNSTINHTRSSVEGLCEGICAGLVDAIWWGTCDYKPKGDASTGWSVSRILIEE 240		
Db	181	GSPENNSTINHTRSSVEGLCEGICAGLVDAIWWGTCDYKPKGDASTGWSVSRILIEE 240		
Qy	241	LPK 243		
Db	241	LPK 243		

1026

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2006, 09:50:45 ; Search time 92.1176 Seconds
(without alignments)
1102.205 Million cell updates/sec

Title: US-10-634-108-4
Perfect score: 1303
Sequence: 1 MRPOGPAPSPQRLRGLLLLL.....GDASTGWSVSRIIIEELPK 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1303	100.0	243	3	US-09-829-472A-19
2	1303	100.0	243	4	US-10-045-992-4
3	1303	100.0	243	4	US-10-097-340-135
4	1303	100.0	243	4	US-10-295-027-50
5	1303	100.0	243	4	US-10-295-027-1173
6	1303	100.0	243	4	US-10-173-999-74
7	1303	100.0	243	4	US-10-058-270A-38
8	1303	100.0	243	4	US-10-634-108-4
9	1303	100.0	243	4	US-10-188-832-175
10	1303	100.0	243	5	US-10-939-233-4
11	1303	100.0	243	6	US-11-050-926-135
12	1303	100.0	278	4	US-10-097-340-131
13	1303	100.0	278	4	US-10-177-293-496
14	1303	100.0	278	4	US-10-301-822-77
15	1303	100.0	278	4	US-10-296-115-1261
16	1303	100.0	278	5	US-10-961-139-2
17	1303	100.0	278	6	US-11-050-926-131
18	1300	99.8	243	4	US-10-351-334-205
19	1300	99.8	278	4	US-10-060-036-4551
20	1298	99.6	243	4	US-10-176-847-14
21	1294	99.3	243	3	US-09-834-759-514
22	1294	99.3	243	3	US-09-938-418-7
23	1294	99.3	243	3	US-09-946-374-352
24	1294	99.3	243	3	US-09-946-374-431
25	1294	99.3	243	4	US-10-006-867-122
26	1294	99.3	243	4	US-10-052-586-366
27	1294	99.3	243	4	US-10-007-803-514

28	1294	99.3	243	4	US-10-063-547-122	Sequence 122, App
29	1294	99.3	243	4	US-10-063-551-122	Sequence 122, App
30	1294	99.3	243	4	US-10-174-590-366	Sequence 366, App
31	1294	99.3	243	4	US-10-176-758-366	Sequence 366, App
32	1294	99.3	243	4	US-10-175-737-366	Sequence 122, App
33	1294	99.3	243	4	US-10-063-616-122	Sequence 366, App
34	1294	99.3	243	4	US-10-174-581-366	Sequence 122, App
35	1294	99.3	243	4	US-10-176-483-366	Sequence 366, App
36	1294	99.3	243	4	US-10-176-749-366	Sequence 366, App
37	1294	99.3	243	4	US-10-176-914-366	Sequence 366, App
38	1294	99.3	243	4	US-10-176-915-366	Sequence 366, App
39	1294	99.3	243	4	US-10-063-569-122	Sequence 122, App
40	1294	99.3	243	4	US-10-063-513-122	Sequence 122, App
41	1294	99.3	243	4	US-10-063-515-122	Sequence 122, App
42	1294	99.3	243	4	US-10-063-512-122	Sequence 122, App
43	1294	99.3	243	4	US-10-173-706-366	Sequence 366, App
44	1294	99.3	243	4	US-10-175-738-366	Sequence 366, App
45	1294	99.3	243	4	US-10-175-752-366	Sequence 366, App

ALIGNMENTS

RESULT 1
US-09-829-472A-19
; Sequence 19, Application US/09829472A
; Publication No. US20040146862A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Breast Cancer
; FILE REFERENCE: 018501-001200US
; CURRENT APPLICATION NUMBER: US/09/829,472A
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 09/525,361
; PRIOR FILING DATE: 2000-03-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO:19
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: BCN4, ESTs, secreted protein
US-09-829-472A-19

Query Match	100.0%	Score 1303	DB 3	Length 243
Best Local Similarity	100.0%	Pred. No. 1.6e-115	Mismatches 0	Indels 0
Matches 243	Conservative 0			Gaps 0
QY	1	MRPOGPAPSPQRLRGLLLLLLQLPAPSSASIPKKGKQAKLRQREVVDLYNGMCIQGPA	60	
Db	1	MRPOGPAPSPQRLRGLLLLLLQLPAPSSASIPKKGKQAKLRQREVVDLYNGMCIQGPA	60	
QY	61	GVPRDGPSPGANGIPGTGPIPORDGFKGKGKGLRSPFESWTPNTYKQCSWSLNYGIDL	120	
Db	61	GVPRDGPSPGANGIPGTGPIPORDGFKGKGKGLRSPFESWTPNTYKQCSWSLNYGIDL	120	
QY	121	GKIACTFTTKMESNLSALRVLFSGSLRLKCRNACCORWYFTFNGAECSGPLPEAIYLDQ	180	
Db	121	GKIACTFTTKMESNLSALRVLFSGSLRLKCRNACCORWYFTFNGAECSGPLPEAIYLDQ	180	
QY	181	GSPENNSTINIHRSTSVGLCEGIGAGLVDAIWTGTCSDYPKGDASTGWSVSRIIIEE	240	
Db	181	GSPENNSTINIHRSTSVGLCEGIGAGLVDAIWTGTCSDYPKGDASTGWSVSRIIIEE	240	
QY	241	LPK 243		
Db	241	LPK 243		

```
RESULT 2
US-10-045-992-4
; Sequence 4, Application US/10045992
; Publication No. US20020161211A1
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhard
; APPLICANT: FRIESEL, Robert E.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODELIN
; FILE REFERENCE: 053689-5006-01
; CURRENT APPLICATION NUMBER: US/10/045,992
; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 09/692,081
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-045-992-4

Query Match 100.0%; Score 1303; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.6e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRPGPAASPORLRLGLLLLLLLQLPAPSSASEIPKGKQKAOIROREVVDLYNGMCLQGPA 60
Db 1 MRPGPAASPORLRLGLLLLLLLQLPAPSSASEIPKGKQKAOIROREVVDLYNGMCLQGPA 60

Qy 61 GVPGRDGGSPGANGIPGTGIPGRDGFKEGKECLRESFEESWTNPKQCSWSSLYNGIDL 120
Db 61 GVPGRDGGSPGANGIPGTGIPGRDGFKEGKECLRESFEESWTNPKQCSWSSLYNGIDL 120

Qy 121 GKIAECTFTKMRNSALRVLPFGSLRLKCRNACCORWYFTTNGAECGSLPIEAIYYLDQ 180
Db 121 GKIAECTFTKMRNSALRVLPFGSLRLKCRNACCORWYFTTNGAECGSLPIEAIYYLDQ 180

Qy 181 GSPENSTINIHRTSSVEGLCEGAGLVDAIWWGTCTSDYPKGDASTGWSVSRIIIEE 240
Db 181 GSPENSTINIHRTSSVEGLCEGAGLVDAIWWGTCTSDYPKGDASTGWSVSRIIIEE 240

Qy 241 LPK 243
Db 241 LPK 243

RESULT 3
US-10-097-340-135
; Sequence 135, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025

; Sequence 50, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464

; Sequence 50, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
```

```
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-50

Query Match      100.0%; Score 1303; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.6e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRQGPAAAPQRLGRLGLLLLLQLPAPSSASBIPKGKQKAQRLQREVVLDYNGMCLQGPA 60
DB 1 MRQGPAAAPQRLGRLGLLLLLQLPAPSSASBIPKGKQKAQRLQREVVLDYNGMCLQGPA 60
QY 61 GVPGRDGGPGANGIPCTPGIPGRDGFKGEGKCLRESPEESWTPNYKQCSWSLNYGIDL 120
DB 61 GVPGRDGGPGANGIPCTPGIPGRDGFKGEGKCLRESPEESWTPNYKQCSWSLNYGIDL 120
QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIYLDQ 180
DB 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIYLDQ 180
QY 181 GSPENNSTINIHRSTSVEGLCEGIGAGLVDAIWIWGTCSDDYPKGDASTGWNVSRIIEE 240
DB 181 GSPENNSTINIHRSTSVEGLCEGIGAGLVDAIWIWGTCSDDYPKGDASTGWNVSRIIEE 240
QY 241 LPK 243
DB 241 LPK 243

RESULT 5
US-10-295-027-1173
; Sequence 1173, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
```

```
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1173
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1173

Query Match      100.0%; Score 1303; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.6e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRQGPAAAPQRLGRLGLLLLLQLPAPSSASBIPKGKQKAQRLQREVVLDYNGMCLQGPA 60
DB 1 MRQGPAAAPQRLGRLGLLLLLQLPAPSSASBIPKGKQKAQRLQREVVLDYNGMCLQGPA 60
QY 61 GVPGRDGGPGANGIPCTPGIPGRDGFKGEGKCLRESPEESWTPNYKQCSWSLNYGIDL 120
DB 61 GVPGRDGGPGANGIPCTPGIPGRDGFKGEGKCLRESPEESWTPNYKQCSWSLNYGIDL 120
QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIYLDQ 180
DB 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIYLDQ 180
QY 181 GSPENNSTINIHRSTSVEGLCEGIGAGLVDAIWIWGTCSDDYPKGDASTGWNVSRIIEE 240
DB 181 GSPENNSTINIHRSTSVEGLCEGIGAGLVDAIWIWGTCSDDYPKGDASTGWNVSRIIEE 240
QY 241 LPK 243
DB 241 LPK 243

RESULT 6
US-10-173-999-74
; Sequence 74, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
```

```
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-74

Query Match      100.0%; Score 1303; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.6e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRPGPAAAPQRLRGLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
Db 1 MRPGPAAAPQRLRGLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60

Qy 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGECLRESFEESWTPNYKQCSWSSLYGIDL 120
Db 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGECLRESFEESWTPNYKQCSWSSLYGIDL 120

Qy 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECSGPLPIEAIYLDQ 180
Db 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECSGPLPIEAIYLDQ 180

Qy 181 GSPENNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWSVSRIIIEE 240
Db 181 GSPENNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWSVSRIIIEE 240

Qy 241 LPK 243
Db 241 LPK 243

RESULT 7
US-10-058-270A-38
; Sequence 38, Application US/10058270A
; Publication No. US20040029114A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Afar, Daniel
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Breast Cancer
; FILE REFERENCE: 016501-005210US
; CURRENT APPLICATION NUMBER: US/10/058,270A
; CURRENT FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: US 60/263,965
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: US 60/265,928
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: US 09/829,472
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 60/282,698
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 60/288,590
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,443
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-058-270A-38

Query Match      100.0%; Score 1303; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.6e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRPGPAAAPQRLRGLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
Db 1 MRPGPAAAPQRLRGLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60

Qy 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGECLRESFEESWTPNYKQCSWSSLYGIDL 120
Db 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGECLRESFEESWTPNYKQCSWSSLYGIDL 120

Qy 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECSGPLPIEAIYLDQ 180
Db 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECSGPLPIEAIYLDQ 180

Qy 181 GSPENNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWSVSRIIIEE 240
Db 181 GSPENNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWSVSRIIIEE 240

Qy 241 LPK 243
Db 241 LPK 243

RESULT 8
US-10-634-108-4
; Sequence 4, Application US/10634108
; Publication No. US20040063176A1
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhard
; APPLICANT: FRIESEL, Robert P.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/10/634,108
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/692,081
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-634-108-4

Query Match      100.0%; Score 1303; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.6e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRPGPAAAPQRLRGLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60
Db 1 MRPGPAAAPQRLRGLLLLLQLPAPSSASEIPKGKQKQALRQREVVDLYNGMCLQGPA 60

Qy 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGECLRESFEESWTPNYKQCSWSSLYGIDL 120
Db 61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGECLRESFEESWTPNYKQCSWSSLYGIDL 120

Qy 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECSGPLPIEAIYLDQ 180
Db 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECSGPLPIEAIYLDQ 180

Qy 181 GSPENNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWSVSRIIIEE 240
Db 181 GSPENNSTINIHRTSSVEGLCEGAGLVDAIIVWGTCSDPKGDASTGWSVSRIIIEE 240

Qy 241 LPK 243
Db 241 LPK 243

RESULT 9
US-10-188-832-175
; Sequence 175, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
```

QY	1	MRPQGPAA	SPQRLG	LLLLLL	LLQLPAP	SSASBI	PKGKQ	KAQLRQ	REVVDL	YNGMCI	QGPA	60
Db	1	MRPQGPAA <td>SPQRLG<td>LLLLLL<td>LLQLPAP<td>SSASBI<td>PKGKQ<td>KAQLRQ<td>REVVDL<td>YNGMCI<td>QGPA</td><td>60</td></td></td></td></td></td></td></td></td>	SPQRLG <td>LLLLLL<td>LLQLPAP<td>SSASBI<td>PKGKQ<td>KAQLRQ<td>REVVDL<td>YNGMCI<td>QGPA</td><td>60</td></td></td></td></td></td></td></td>	LLLLLL <td>LLQLPAP<td>SSASBI<td>PKGKQ<td>KAQLRQ<td>REVVDL<td>YNGMCI<td>QGPA</td><td>60</td></td></td></td></td></td></td>	LLQLPAP <td>SSASBI<td>PKGKQ<td>KAQLRQ<td>REVVDL<td>YNGMCI<td>QGPA</td><td>60</td></td></td></td></td></td>	SSASBI <td>PKGKQ<td>KAQLRQ<td>REVVDL<td>YNGMCI<td>QGPA</td><td>60</td></td></td></td></td>	PKGKQ <td>KAQLRQ<td>REVVDL<td>YNGMCI<td>QGPA</td><td>60</td></td></td></td>	KAQLRQ <td>REVVDL<td>YNGMCI<td>QGPA</td><td>60</td></td></td>	REVVDL <td>YNGMCI<td>QGPA</td><td>60</td></td>	YNGMCI <td>QGPA</td> <td>60</td>	QGPA	60
QY	61	GVPGRD <td>SGPGANG</td> <td>IPGT</td> <td>PGI</td> <td>PGRD</td> <td>GPKG</td> <td>KGCL</td> <td>RES</td> <td>FEES</td> <td>WT</td> <td>PNYKQCSWS</td>	SGPGANG	IPGT	PGI	PGRD	GPKG	KGCL	RES	FEES	WT	PNYKQCSWS
Db	61	GVPGRD <td>SGPGANG</td> <td>IPGT</td> <td>PGI</td> <td>PGRD</td> <td>GPKG</td> <td>KGCL</td> <td>RES</td> <td>FEES</td> <td>WT</td> <td>PNYKQCSWS</td>	SGPGANG	IPGT	PGI	PGRD	GPKG	KGCL	RES	FEES	WT	PNYKQCSWS
QY	121	GKIAECT	TKMRS	NALRV	LSGSL	RKCRNA	ACCQ	RWY	TF	FN	GABCSG	PLPIE
Db	121	GKIAECT	TKMRS	NALRV	LSGSL	RKCRNA	ACCQ	RWY	TF	FN	GABCSG	PLPIE
QY	181	GSPEN	NTIN	IHR	TSS	VGL	CEG	IGAG	LVD	VAI	WVGT	CSDY
Db	181	GSPEN	NTIN	IHR	TSS	VGL	CEG	IGAG	LVD	VAI	WVGT	CSDY
QY	241	LPK	243									
Db	241	LPK	243									
RESULT 11												
US-11-050-926-135												
; Sequence 135, Application US/11050926												
; Publication No. US20050214831A1												
; GENERAL INFORMATION:												
; APPLICANT: John MONAHAN												
; APPLICANT: Manjula GANNAVAPU												
; APPLICANT: Sebastian HOERSCH												
; APPLICANT: Shubhangi KAMATKAR												
; APPLICANT: Steve G. KOVATS												
; APPLICANT: Rachel E. MEYERS												
; APPLICANT: Michael MORRISEY												
; APPLICANT: Peter OLANDT												
; APPLICANT: Ami SEN												
; APPLICANT: Peter VEIBY												
; APPLICANT: Gordon B. MILLS												
; APPLICANT: Robert C. BAST, Jr.												
; APPLICANT: Karen LU												
; APPLICANT: Rosemarie SCHMANDT												
; APPLICANT: Xumei ZHAO												
; APPLICANT: Karen GLATT												
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification												
; FILE REFERENCE: MRI-030												
; CURRENT APPLICATION NUMBER: US/11/050,926												
; PRIOR FILING DATE: 2005-02-04												
; PRIOR APPLICATION NUMBER: US/10/097,340												
; PRIOR FILING DATE: 2002-03-14												
; PRIOR APPLICATION NUMBER: 60/276,025												
; PRIOR FILING DATE: 2001-03-14												
; PRIOR APPLICATION NUMBER: 60/325,149												
; PRIOR FILING DATE: 2001-09-26												
; PRIOR APPLICATION NUMBER: 60/276,026												
; PRIOR FILING DATE: 2001-03-14												
; PRIOR APPLICATION NUMBER: 60/324,967												
; PRIOR FILING DATE: 2001/09/26												
; PRIOR APPLICATION NUMBER: 60/311,732												
; PRIOR FILING DATE: 2001-08-10												
; PRIOR APPLICATION NUMBER: 60/325,102												
; PRIOR FILING DATE: 2001-09-26												
; PRIOR APPLICATION NUMBER: 60/323,580												
; PRIOR FILING DATE: 2001-09-19												
; NUMBER OF SEQ ID NOS: 363												
; SOFTWARE: FastSeq for Windows Version 4.0												
; SEQ ID NO 135												
; LENGTH: 243												
; TYPE: PRT												
; ORGANISM: Homo sapiens												
US-11-050-926-135												
Query Match 100.0%; Score 1303; DB 6; Length 243;												

[illegible]

```

RESULT 12
US-10-097-340-131
; Sequence 131, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAWATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 131
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-131

```

Query Match 100.0%; Score 1303; DB 4; Length 278;

[illegible]

RESULT 13

US-10-177-293-496

Sequence 496, Application US/10177293

Publication No. US20030124128A1

GENERAL INFORMATION:

APPLICANT: Lillie, James

APPLICANT: Glatt, Karen

APPLICANT: Zhao, Xumei

APPLICANT: Gannavarpu, Manjula

APPLICANT: Kamatkar, Shubhangi

APPLICANT: Mertens, Maureen

APPLICANT: Myer, Vic

APPLICANT: Wang, Youzhen

APPLICANT: Xu, Yongyao

APPLICANT: Hoersch, Sebastian

APPLICANT: Monahan, John

APPLICANT: Meyers, Rachel E.

APPLICANT: Baat Jr., Robert C.

APPLICANT: Hortobagyi, Gabriel N.

APPLICANT: Pusztai, Lajos

APPLICANT: Meric, Funda

APPLICANT: Sahin, Aysegul

APPLICANT: Mills, Gordon B.

TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST CANCER

FILE REFERENCE: MRI-038

CURRENT APPLICATION NUMBER: US/10/177,293

PRIOR FILING DATE: 2002-06-21

PRIOR APPLICATION NUMBER: US 60/299,887

PRIOR FILING DATE: 2001-06-21

PRIOR APPLICATION NUMBER: US 60/301,572

PRIOR FILING DATE: 2001-06-27

PRIOR APPLICATION NUMBER: US 60/306,501

PRIOR FILING DATE: 2001-07-18

PRIOR APPLICATION NUMBER: US 60/325,002

PRIOR FILING DATE: 2001-09-25

PRIOR APPLICATION NUMBER: US 60/362,585

PRIOR FILING DATE: 2002-03-05

PRIOR APPLICATION NUMBER: US 60/xxx,xxx

PRIOR FILING DATE: 2002-05-14

NUMBER OF SEQ ID NOS: 506

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 496

LENGTH: 278

TYPE: PRI

ORGANISM: Homo sapiens

US-10-177-293-496

Query Match 100.0%; Score 1303; DB 4; Length 278;

[illegible]

```

RESULT 14
US-10-301-822-77
; Sequence 77, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: BURGART, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MP001-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-101-822-77

```

	Query Match	100.0%;	Score 1303;	DB 4;	Length 278;
	Best Local Similarity	100.0%;	Pred. No. 1.9e-115;		
	Matches 243;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MRPGGPAASQRLRGILLLLQLLPAPSSASEIPKGKQKQALRQREVVVDLYNGMCLGQPA	60		
Db	36	MRPGGPAASQRLRGILLLLQLLPAPSSASEIPKGKQKQALRQREVVVDLYNGMCLGQPA	95		
Qy	61	GVPGRDGSPGANGIPGTPGIPGRDGFKGEGECLRESFEESWTENYKQCSWSSLYNGIDL	120		
Db	96	GVPGRDGSPGANGIPGTPGIPGRDGFKGEGECLRESFEESWTENYKQCSWSSLYNGIDL	155		
Qy	121	GKTAECTFTQWRNSALRVLPFSGSLRLKRNACCQRYFFTPNGAECGSPPIEAIYYLQD	180		
Db	156	GKTAECTFTQWRNSALRVLPFSGSLRLKRNACCQRYFFTPNGAECGSPPIEAIYYLQD	215		

Qy	181	GSPEMNSTINIHRTSSVEGLCEGIGAGLVDAI	WVGTCSDYPKGDASTGWSVSRIIIEE	240
Db	216	GSPEMNSTINIHRTSSVEGLCEGIGAGLVDAI	WVGTCSDYPKGDASTGWSVSRIIIEE	275
Qy	241	LPK 243		
Db	276	LPK 278		

RESULT 15

US-10-296-115-1261

; Sequence 1261, Application US/10296115

; Publication No. US20040053248A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq Inc

; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides

; FILE REFERENCE: 784PCT

; CURRENT APPLICATION NUMBER: US/10/296,115

; CURRENT FILING DATE: 2002-11-18

; PRIOR APPLICATION NUMBER: US09/488,725

; PRIOR FILING DATE: 2000-01-21

; PRIOR APPLICATION NUMBER: US09/552,317

; PRIOR FILING DATE: 2000-04-25

; NUMBER OF SEQ ID NOS: 1478

; SEQ ID NO 1261

; LENGTH: 278

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-296-115-1261

Query Match	100.0%	Score 1303;	DB 4;	Length 278;
Best Local Similarity	100.0%;	Pred. No. 1.9e-115;		
Matches 243;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MRPQGPAAAPQRLRCILLILLILLQLPAPSSASEIPKGKQAKQLRQREVVDLYNGMCLQGPA	60	
Db	36	MRPQGPAAAPQRLRGLLLILLILLQLPAPSSASEIPKGKQAKQLRQREVVDLYNGMCLQGPA	95	
Qy	61	GVPRGDGSPGANGIPGTPGIPGRDGFKGEGCELRSPESSWTPNPKQCSWSSLYNGIDL	120	
Db	96	GVPRGDGSPGANGIPGTPGIPGRDGFKGEGCELRSPESSWTPNPKQCSWSSLYNGIDL	155	
Qy	121	GKTAECTFTKWRNSALRVLPFSGSLRLKCRNACCQWYFTFNGAECSPGLPEAIYYLDQ	180	
Db	156	GKTAECTFTKWRNSALRVLPFSGSLRLKCRNACCQWYFTFNGAECSPGLPEAIYYLDQ	215	
Qy	181	GSPEMNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCDSPYKGDASTGWNVSRIIEE	240	
Db	216	GSPEMNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCDSPYKGDASTGWNVSRIIEE	275	
Qy	241	LPK 243		
Db	276	LPK 278		

Search completed: April 20, 2006, 09:56:57
Job time : 94.1176 secs

THIS PAGE BLANK (USPTO)